University of Alberta

Determinants of PAX3 behavior: A molecular and cellular analysis of the PAX3 transcription factor and disease-associated mutants

by

Gareth Neill Corry ©

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Medical Sciences – Medical Genetics

Edmonton, Alberta
Spring 2008

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-45411-4
Our file Notre référence
ISBN: 978-0-494-45411-4

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

The PAX3 transcription factor is a key regulator of developmentally important processes in metazoan organisms and also carries out distinct postnatal functions. PAX3 is characterized by two DNA-binding modules, the paired domain (PD) and homeodomain (HD), which functionally interact to control target gene expression. Mutations in PAX3 cause Waardenburg syndrome in humans and the mouse *Splotch* phenotype, both of which are characterized by defects in neural crest-derived and myogenic cell types. Through molecular and cellular analyses, we have investigated the determinants that underlie PAX3 behavior and have shed light on how disease-causing mutations affect these determinants. We show that PAX3 uses different arrangements of its DNA-binding domains to achieve distinct modes of target sequence recognition. Notably, disease mutations exert varying effects on PAX3 target affinity, specificity, and functional cooperativity between the PD and HD. Additionally, we find that these mutations exert variable effects, including repression and enhanced expression, on the activation of PAX3-responsive reporter genes. *In vivo* studies show that PAX3 localizes to the interchromatin space and displays limited co-localization with sites of transcriptional activity. Importantly, each PAX3 mutant examined in live cells displayed differences in intranuclear mobility compared to the wild type protein and, based on their intranuclear compartmentalization and dynamics, we were able to separate our cohort of mutations into two distinct classes. Significantly, this classification method helps clarify the determinants that regulate PAX3 behavior *in vivo* and has enabled us to propose a model in which PAX3 activity is regulated by a two-step process involving compartmentalization to a

particular nuclear domain, followed by identification and interaction with chromatin-embedded regulatory target sequences. Our results also suggest that conformational rearrangements during each step play a role in optimizing PAX3 performance. Together, our results provide important information regarding the determinants that regulate PAX3 behavior and the associated effects of disease mutations, and establish a foundation for integrating molecular and cellular analyses of transcription factor function.

ACKNOWLEDGEMENTS

Immediate thanks go to Dr. Alan Underhill for taking a chance way back in the summer of '99 and letting me do some part-time work in the lab, then keeping me on as a project student, research assistant, and finally, a graduate student. I think everything I know about scientific research is due in large part to Alan's guidance, advice, and instruction and I credit my research accomplishments to his supervision and assistance. I wish him every success for the future.

Thanks to Dr. Michael Walter and Dr. Mark Glover for contributing their time to be part of my graduate committee and helping to direct my research; their feedback and suggestions over the years have been much appreciated. I also thank Dr. Jim Davie, Dr. Roseline Godbout, and Dr. Stacey Bléoo for being part of my examining committee, and Dr. Rachel Wevrick, Dr. Michael Schultz, and Dr. Moira Glerum for participating in my PhD candidacy exam.

My research would not have progressed without the help of other students, postdoctoral fellows, technicians, and professors. In particular, I express thanks to Ning Hu for her technical assistance in the lab and for taking care of our tissue culture. I also appreciate the support and helpfulness of current and former members of the Underhill lab during my studies. Additional thanks go to the other Medical Genetics labs, as well as the labs of Drs. Michael Schultz, Michael Hendzel, John Greer, and Shairaz Baksh for sharing reagents, advice, and expertise. Special thanks to Jason Bush for providing primary cell cultures, Robin Clugston for providing rat livers, and Kristal Missiaen for setting aside her own work to do our FRAP experiments for us. Thanks also to Dr. Xuejun Sun and Gerry Barron at the Cross Cancer Institute microscope facility for their assistance and patience, and Dr. Colin Goding, Dr. Frederic Barr, and Dr. Peter Adams for providing reagents.

I acknowledge the various agencies and institutions that have provided funding during my studies, including the University of Alberta, the Natural Sciences and Engineering Research Council of Canada, the Canadian Institutes of Health Research, and Alberta Learning.

Finally, and simply, thank you to my parents and family for their continual support, advice, understanding, and love – past, present, and future.

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION 1

1.1 PAX3: structure, function, and disease 2

- **1.1.1** PAX proteins 2
- **1.1.2** *Pax3/PAX3* gene structure and regulatory elements 3
- **1.1.3** PAX3 protein structure 7
- **1.1.4** *Pax3* expression during development 9
- **1.1.5** PAX3-related diseases 12
- **1.1.6** Role of PAX3 in developmental pathways 16
- **1.1.7** PAX3 target genes 19
- **1.1.8** PAX3 protein-protein interactions 25
- **1.1.9** Functional interactions involving the paired domain and homeodomain 27

1.2 PAX3: from the test tube to the nucleus 29

1.3 Intranuclear organization of transcription 30

- **1.3.1** The eukaryotic nucleus 30
- **1.3.2** Subnuclear localization of transcription factors 35
- **1.3.3** Intrinsic determinants of transcription factor subnuclear localization 39
- **1.3.4** Regulation of transcription factor subnuclear localization through co-factor interactions 42
- **1.3.5** Intranuclear dynamics 44
- **1.3.6** Functional organization of eukaryotic transcription 48

1.4 Objective 51

CHAPTER 2. METHODS AND MATERIALS 53

2.1 Plasmid construction 54

- **2.1.1** Bacterial expression constructs 54
- **2.1.2** Mammalian expression constructs 61
- **2.1.3** Luciferase reporter constructs 62

2.2 Mutant expression constructs....64
2.3 Hexahistidine-tagged protein expression and purification....64
2.4 Electrophoretic mobility shift assays....67
2.5 Bulk chromatin isolation....70
2.6 Affinity chromatography pull-down assays....70
2.7 Cell lines and culture....72
2.8 Dual luciferase assays....72
2.9 Antibodies....73
2.10 Whole cell extracts....73
2.11 Western transfer....74
2.12 Immunocytochemistry....74
2.13 Fluorescence recovery after photobleaching....75
2.14 Bioinformatics....76

CHAPTER 3. PAX3-TARGET INTERACTION....77
3.1 Background....78
3.2 Analysis of PAX3 recognition sequences....79
3.3 Evolutionary conservation of selected PAX3 target sequences....83
3.4 DNA-binding characteristics of PAX3....86
3.5 The PAX3 homeodomain increases binding affinity for composite sequences....89
3.6 Analysis of PAX3 interaction with the *Trp-1* promoter sequence....91
3.7 Sequence determinants for *MITF* promoter binding by PAX3....93
3.8 Sequence determinants within the *MITF* hd3 motif....95
3.9 Structural requirements for PAX3-*MITF* interaction....98
3.10 Participation of paired domain and homeodomain homologs in cooperative binding....101
3.11 Effects of disease mutations on PAX3 target recognition....105
3.12 Effects of PAX3 disease mutations on paired domain and homeodomain cooperativity....109

3.13 Functional analysis of the *MITF* and *Trp-1* PAX3 recognition elements 111
3.14 Discussion 113

CHAPTER 4. CHARACTERIZATION OF PAX3 IN THE NUCLEUS 118
4.1 Background 119
4.2 Subnuclear localization of PAX3 121
4.3 Localization of PAX3 with respect to post-translational histone modifications 124
4.4 Co-localization of PAX3 and sites of transcriptional activity 128
4.5 PAX3 DNA-binding domains constitute separable determinants of subnuclear localization 128
4.6 Effect of flanking regions on PAX3 homeodomain subnuclear localization 134
4.7 Effects of disease mutations on PAX3 subnuclear localization 136
4.8 Nuclear dynamics of wild type and mutant PAX3 139
4.9 Discussion 145

CHAPTER 5. DETERMINANTS OF PAIRED-TYPE HOMEODOMAIN BEHAVIOR 152
5.1 Background 153
5.2 Comparison of the PAX3, PRRX1, and PITX2 homeodomains 155
5.3 Subnuclear localization of the PAX3, PRRX1, and PITX2 homeodomains 157
5.4 Analysis of paired-type homeodomain target preference 159
5.5 Mutation of residue 50 alters HD intranuclear compartmentalization 162
5.6 Intranuclear mobility of wild type and residue 50-mutant Paired-type homeodomains 165
5.7 Structural compatibility of homeodomain residues influences DNA binding and subnuclear localization 167
5.8 The PAX3 homeodomain exhibits unique chromatin interaction properties 172
5.9 Discussion 174

CHAPTER 6. CONCLUSIONS & OUTLOOK 180
6.1 Summary 181
6.2 A two-step model for regulation of PAX3 function *in vivo* 182
6.3 Do distinct conformations underlie PAX3 behavior and function? 185
6.4 PAX3: activator *and* repressor? 189
6.5 PAX3 and Waardenburg syndrome: towards a genotype-phenotype correlation 193
6.6 Outlook 194

BIBLIOGRAPHY 198

APPENDIX 1. EFFECT OF THE INTERDOMAIN LINKER ON PAX3 DNA BINDING 238
A1.1 Role of the interdomain linker in PAX3 homeodomain DNA binding 239
A1.2 Influence of the linker on cooperative binding 242
A1.3 Discussion 245

APPENDIX 2. RELATIONSHIP BETWEEN PAX3 AND THE TLE4 CO-REPRESSOR 247
A2.1 TLE4: gene and protein structure 248
A2.2 Intranuclear relationship between PAX3 and TLE4 250
A2.3 Disease mutations affect PAX3 co-localization with TLE4 isoforms 252
A2.4 Discussion 255

LIST OF TABLES

CHAPTER 2

Table 1. Primers used for PCR amplification of inserts for expression constructs 55
Table 2. PAX3 expression constructs in pET21a 59
Table 3. Recombinant paired domain and homeodomain expression constructs in pET21a....60
Table 4. PAX3 expression constructs in pEGFP-N1 63
Table 5. Primers for site-directed mutagenesis....65
Table 6. Oligonucleotides used for electrophoretic mobility shift assays 68

CHAPTER 5

Table 7. Target preference of wild type, residue 50-mutant, and chimeric homeodomains 170

LIST OF FIGURES

CHAPTER 1

Figure 1-1. Alternative splicing events create several *PAX3* transcripts 4

Figure 1-2. Mammalian PAX proteins 6

Figure 1-3. Developmental pathways involving PAX3 10

Figure 1-4. Positions of *PAX3* lesions associated with Waardenburg syndrome 14

Figure 1-5. Organization of eukaryotic gene expression 50

CHAPTER 2

Figure 2-1. Affinity chromatography pull-down assays and analysis of PAX3- or Paired-type homeodomain interaction with chromatin and associated nucleic acid 71

CHAPTER 3

Figure 3-1. Alignment of selected PAX3 target recognition elements 80

Figure 3-2. Evolutionary conservation of PAX3 target promoter sequences 84

Figure 3-3. Interaction of PAX3 with selected target promoter sequences 87

Figure 3-4. The homeodomain contributes to PAX3 interaction with composite target elements 90

Figure 3-5. Analysis of PAX3 binding to the *Trp-1* promoter 92

Figure 3-6. Both the PAX3 paired domain and homeodomain make important contributions to *MITF* promoter binding 94

Figure 3-7. Analysis of the *MITF* hd3 motif 97

Figure 3-8. Composite promoter elements support cooperative binding by the PAX3 paired domain and homeodomain 99

Figure 3-9. Paired domain paralogs can recruit the PAX3 homeodomain to the *MITF* promoter sequence 102

Figure 3-10. PAX3/7 homeodomain homologs cannot cooperatively bind to a composite sequence with the PAX3 paired domain 104

Figure 3-11. Effects of disease mutations on PAX3 DNA binding behavior........106
Figure 3-12. PAX3 disease mutations interfere with inter-domain cooperativity........110
Figure 3-13. Functional analysis of the *MITF* and *Trp-1* PAX3-binding sites........112

CHAPTER 4

Figure 4-1. Subnuclear localization of PAX3........122
Figure 4-2. Distribution of PAX3 relative to post-translational histone modifications........126
Figure 4-3. PAX3 shows limited overlap with sites of transcriptional activity........129
Figure 4-4. The PAX3 paired domain and homeodomain represent separable determinants of intranuclear localization........131
Figure 4-5. Tethering the PAX3 DNA-binding domains constrains their mobility in live cells........133
Figure 4-6. Flanking regions influence the subnuclear localization of the PAX3 homeodomain........135
Figure 4-7. Disease mutations alter the subnuclear localization of PAX3........137
Figure 4-8. Behavior of wild type and mutant PAX3-GFP in live cells........140
Figure 4-9. Mutations causing compartmentalization defects exert a dominant effect on PAX3 behavior in the nucleus........144
Figure 4-10. Disease mutations exert differential effects on determinants that constrain PAX3 activity in the nucleus........150

CHAPTER 5

Figure 5-1. Comparison of the PAX3, PRRX1, and PITX2 homeodomains........156
Figure 5-2. Comparison of the subnuclear localization of the PAX3, PRRX1, and PITX2 homeodomains........158
Figure 5-3. Homeodomains from the PAX3, PRRX1, and PITX2 proteins display overlapping target specificity........160

Figure 5-4. Mutation of residue 50 affects homeodomain subnuclear localization 163
Figure 5-5. Intranuclear dynamics of wild type and residue 50-mutant homeodomains 166
Figure 5-6. Contribution of regions outside the recognition helix to homeodomain behavior 168
Figure 5-7. The PAX3 homeodomain possesses unique chromatin-interaction determinants 173

CHAPTER 6

Figure 6-1. A two-step model for regulation of PAX3 function *in vivo* 184
Figure 6-2. A putative composite sequence-binding conformation may suppress an autoinhibitory region in the PAX3 linker 187

APPENDIX 1

Figure A1-1. The PAX3 interdomain linker influences homeodomain DNA binding behavior 240
Figure A1-2. The interdomain linker affects cooperative binding by the paired domain and homeodomain 243

APPENDIX 2

Figure A2-1. The human *TLE4* transcript is alternatively spliced to produce three isoforms 249
Figure A2-2. Intranuclear relationship between PAX3 and TLE4 251
Figure A2-3. Co-localization of TLE4b and TLE4c with selected PAX3 disease mutants 254

ABBREVIATIONS

5-FUrd	5-fluorouridine
aa	amino acids
AcH3	acetylated histone H3
AcH4	acetylated histone H4
ADP	adenosine diphosphate
ALR	acidic linker region
AML	acute myeloid leukemia
AR	androgen receptor
ATP	adenosine triphosphate
Bcd	bicoid
BMP	bone morphogenetic protein
bp	base pair
BrdU	bromodeoxyuridine
C/EBPα	CCAAT/enhancer binding protein α
CBP	CREB-binding protein
CDHS	craniofacial-deafness-hand syndrome
ChIP	chromatin immunoprecipitation
DAPI	4',6-diamidino-2-phenylindole
DAXX	death-associated protein 6
Dct	*dopachrome tautomerase*
Dl	dorsal
DMEM	Dulbecco's Modified Eagle's Medium

DNase	deoxyribnuclease
dpc	days post coitum
DSHB	Developmental Studies Hybridoma Bank
EFT	elongation factor
Eh-1	engrailed homology motif-1
En	engrailed
ER	estrogen receptor
ETO	eight-twenty-one
ETS1	v-ets erythroblastosis virus E26 oncogene homolog 1
eve	*even-skipped*
FRAP	fluorescence recovery after photobleaching
Ftz	fushi tarazu
GFP	green fluorescent protein
GH	growth hormone
GR	glucocorticoid receptor
GRIP-1	glucocorticoid receptor-interacting protein 1
Gro	groucho
gsb	*gooseberry*
H3K4me3	trimethyl-lysine 4 of histone H3
H3K9me3	trimethyl-lysine 9 of histone H3
H3K36me3	trimethyl-lysine 36 of histone H3
H4K20me3	trimethyl-lysine 20 of histone H4
HA	hemagglutinin

HD	homeodomain
HDAC	histone deacetylase
HGF/SF	hepatocyte growth factor/scatter factor
HIC1	hypermethylated in cancer 1
HIRA	HIR histone cell cycle regulation defective homolog A
HMG	high-mobility-group
HP1	heterochromatin protein 1
HSF1	heat shock transcription factor 1
HTH	helix-turn-helix
KAP1	KRAB-associated protein 1
kb	kilobase
kDa	kilodaltons
LEF	lymphoid enhancer-binding factor
Lz	lozenge
MAD	matrix-associated deacetylase
Mb	megabase
MBP	*myelin basic protein*
MEIS	myeloid ecotropic viral integration site
Mitf	*microphthalmia-associated transcription factor*
MMTV	mouse mammary tumor virus
MNase	micrococcal nuclease
MR	mineralocorticoid receptor
MSE	melanocyte-specific initiator element

MSF	melanocyte specific factor
MSX1	msh homeobox 1
N-CAM	*neural cell adhesion molecule*
N-CoR	nuclear receptor co-repressor
Ng-CAM	*neuron-glia cell adhesion molecule*
NLS	nuclear localization signal
NMTS	nuclear matrix-targeting sequence
NTD	neural tube defect
OPT	OCT1/PTF/transcription
PAGE	polyacrylamide gel electrophoresis
PBX	pre-B-cell leukemia homeobox
PD	paired domain
PEI	polyethylenimine
PIAS1	protein inhibitor of activated STAT 1
PITX2	paired-like homeodomain 2
PML	promyelocytic leukemia
POU	PIT-1/OCT/Unc-86
POU_{HD}	POU homeodomain
POU_{S}	POU-specific domain
pRB	retinoblastoma
prd	*paired*
PRRX1	paired related homeobox 1
PRS	paired domain recognition site

VDR	vitamin D receptor
WS	Waardenburg syndrome
WT1	Wilms' Tumor transcription factor
YAP	yes-associated protein
YY1	yin-yang-1

CHAPTER 1. INTRODUCTION

Portions of this chapter are published in:

Corry GN, Underhill DA. (2008) Pax3 and Waardenburg Syndrome Type 1. In, *Inborn Errors of Development, 2nd ed.*, Eds., Epstein CJ, Erickson, RP, and Wynshaw-Boris, A. Oxford, Oxford University Press (*in press*).

Corry GN, Underhill DA. (2005) Subnuclear compartmentalization of sequence-specific transcription factors and regulation of eukaryotic gene expression. *Biochem. Cell Biol.* **83**: 535-47.

Fig. 1-2 is adapted from Underhill DA. (2000) *Biochem. Cell Biol.* **78**: 629-38.

1.1 PAX3: structure, function, and disease

1.1.1 PAX proteins

With the cloning of the *Drosophila paired* (*prd*) gene (Kilchherr et al. 1986) and subsequent characterization of its distinguishing feature, the paired domain (PD) (Bopp et al. 1986), a novel class of DNA-binding proteins was introduced. The *prd* mutant displays an absence of adjacent segmental portions with a two-segment periodicity (Nusslein-Volhard and Wieschaus 1980; Kilchherr et al. 1986), in accordance with the function of *prd* as a pair-rule gene. Based on the discovery of a set of genes that contained similar domains, *prd* was originally hypothesized to be part of a genetic network whose members regulated specific processes during development (Bopp et al. 1986; Frigerio et al. 1986). It is now known that these proteins, defined by the 128-amino acid (aa) PD, function as sequence-specific DNA-binding transcription factors during early fly development. The paired box motif has since been found in evolutionarily diverse organisms (Burri et al. 1989; Balczarek et al. 1997), and the study of PD factor mutants (see Section 1.1.5) confirms their role as important functional determinants of metazoan development. PD factors identified in mice and humans are termed the Paired box (PAX) proteins (Walther et al. 1991; Stuart et al. 1994; Tremblay and Gruss 1994). The following section introduces one of these proteins, PAX3, including its gene and protein structure and involvement in regulating target gene expression and developmental processes. Also discussed are the effects of disease mutations in the murine *Pax3* and human *PAX3* genes and their associated phenotypes and disorders.

1.1.2 *Pax3/PAX3* gene structure and regulatory elements

During a screen of a human genomic DNA library, Burri et al. (1989) identified three clones featuring paired box sequences. Two clones, termed HuP1 and HuP2, showed considerable homology to each other and to *Drosophila prd-gooseberry* (*gsb*) class PD proteins. With the publication of the mouse *Pax3* sequence, HuP2 was shown to correspond to the human ortholog of *Pax3* (Goulding et al. 1991). The mouse *Pax3* locus is situated on chromosome 1 (Evans et al. 1988) and its open reading frame covers 1,437 nucleotides and contains ten exons (Goulding et al. 1991). The human *PAX3* locus maps to chromosome 2q35 (Tsukamoto et al. 1992) and, while initial characterization of *PAX3* showed it was encoded by eight exons (Macina et al. 1995), further examination of its genomic structure revealed two additional exons after exon 8 (Barber et al. 1999).

Alternative splicing events create several characterized PAX3 isoforms (Fig. 1-1). The canonical and most abundant of these, *PAX3c*, is encoded by the first eight exons and produces a 479-aa protein with a relative molecular mass of ~56 kilodaltons (kDa) (Goulding et al. 1991). The next most abundant isoform, *PAX3d*, lacks intron 9 but contains intron 10, while splicing of introns 9 and 10 leads to expression of the *PAX3e* isoform (Barber et al. 1999). Two truncated isoforms, *PAX3a* and *PAX3b*, encode the amino-terminal portion of PAX3 to the end of the PD (Tsukamoto et al. 1994) and two other isoforms, *PAX3g* and *PAX3h*, appear to be specific to human melanomas (Parker et al. 2004). Functional characterization of these seven *PAX3* isoforms in mouse melanocytes was recently performed by Wang et al. (2006), who showed that PAX3c and PAX3d had positive effects on cell growth, migration, and survival, while the other isoforms exerted variable effects on proliferation, migration, apoptosis, transformation,

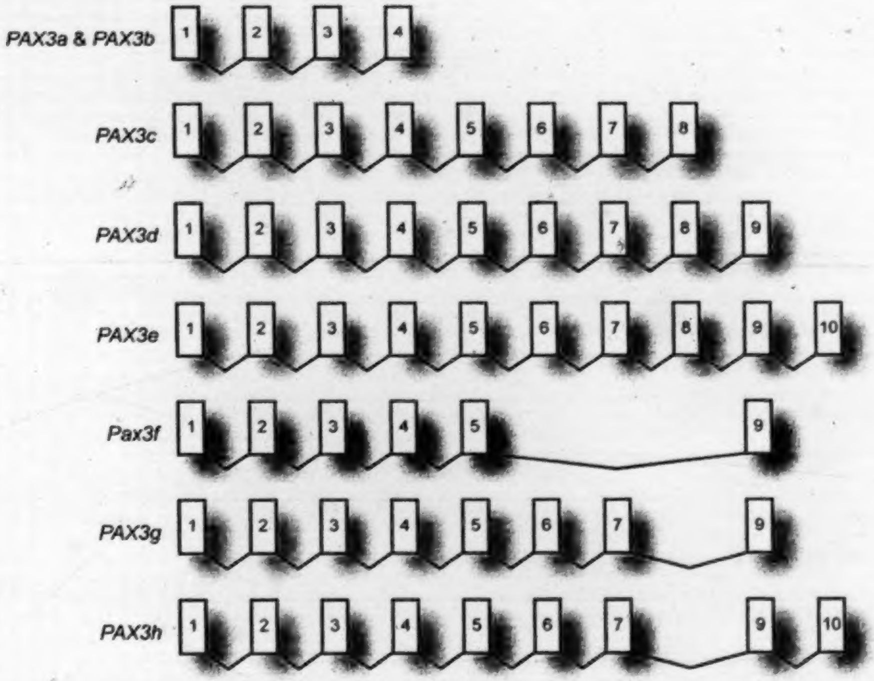
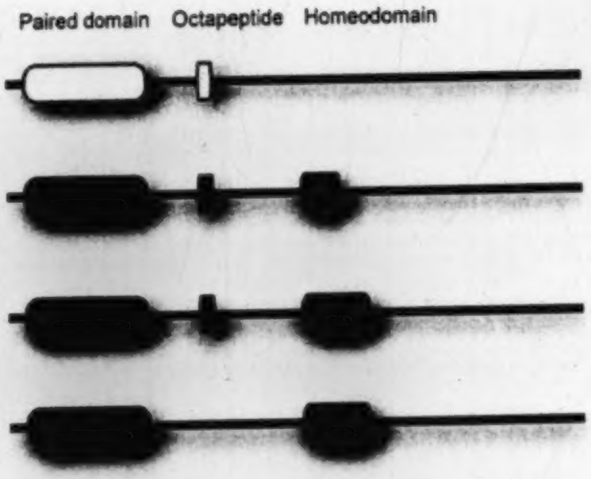

PAX3a & PAX3b
1 2 3 4
PAX3c
1 2 3 4 5 6 7 8
PAX3d
1 2 3 4 5 6 7 8 9
PAX3e
1 2 3 4 5 6 7 8 9 10
Pax3f
1 2 3 4 5 9
PAX3g
1 2 3 4 5 6 7 9
PAX3h
1 2 3 4 5 6 7 9 10

Figure 1-1. Alternative splicing events create several *PAX3* transcripts. *PAX3a* and *PAX3b* are putative isoforms truncated after the paired domain; *PAX3c* is the most commonly described isoform; *PAX3g* and *PAX3h* are enriched in melanomas. Each *PAX3* isoform encodes a unique carboxy-terminal amino acid sequence. The paired domain is encoded by sequences in exons 2, 3, and 4; the homeodomain is encoded by sequences in exons 5 and 6 (see also Fig. 1-4). See Section 1.1.2 for details.

and growth in soft agar. A final putative isoform, *Pax3f*, has been identified in mice and is created by the direct joining of exons 5 and 9 (Barber et al. 1999), but no functional characterization of this protein has been performed to date. Additionally, alternate usage of a splice acceptor between exons 3 and 4 creates versions of the PAX3 isoforms that differ by the inclusion of a single Gln residue at position 108 (Vogan et al. 1996). The inclusion (Q+) or exclusion (Q-) of this Gln influences PAX3 DNA binding activity, particularly in the use of the carboxy-terminal subdomain of the PD (Vogan et al. 1996; Vogan and Gros 1997).

Characterization of the *Pax3* promoter and upstream regulatory region revealed that a ~1.6 kilobase (kb) sequence 5' to the transcription start site was necessary and sufficient for expression of a coupled reporter gene in the dorsal neural tube (Natoli et al. 1997). Accordingly, this genomic segment, together with the PAX3 coding region, was sufficient to rescue neural tube and neural crest-derived defects in PAX3-null mice (Li et al. 1999a); however, an additional region encompassing 14 kb upstream of the promoter conferred expression in migratory myoblast cells and cell populations anterior to the hindbrain and was required for restricting *Pax3* expression from the tail bud. Subsequent analysis of this upstream region revealed discrete enhancer elements, one of which directs PAX3 expression in the hypaxial somite (Brown et al. 2005). Further characterization of the *PAX3* promoter revealed the presence of potential regulatory elements for a number of transcription factors, including AP1, AP2, SP1, TFIID, PEA3, UBP1, EFT, TBP, and MEF2 (Okladnova et al. 1999). Also found in the 5' portion of the human *PAX3* gene was a polymorphic $(CA)_n$ microsatellite, which contained anywhere between 13 and 30 repeats. Together, these studies indicate that the temporal



Group		Chromosome Mouse	Chromosome Human	KO	Mutants Mouse	Mutants Human	Tumor-associated
I	PAX1	2	20p11.2	✓	*Un*		×
	PAX9	12	14q12-q13	✓		Oligodontia	×
II	PAX2	19	10q24.3-q25.1	✓	*Krd/1Neu*	Renal coloboma	✓
	PAX5	4	9p13	✓			✓
	PAX8	2	2q12-q14	✓		Congenital hypothyroidism	✓
III	PAX3	1	2q35		*Sp*	Waardenburg	✓
	PAX7	4	1p36.2-p36.12	✓			✓
IV	PAX4	6	7q32	✓			×
	PAX6	2	11p13		*Sey*	Aniridia	×

Figure 1-2. Mammalian PAX proteins. The nine mammalian PAX proteins are classified into four subgroups based on structural characteristics and sequence homology. All PAX proteins feature a paired domain (PD); class III and IV proteins also possess a homeodomain (HD). Class II proteins contain a truncated HD that lacks the helix-turn-helix motif. Class I, II, and III proteins contain an octapeptide motif in the interdomain linker. Chromosomal locations of corresponding genes are indicated for mouse and human. Characterized mutations of mouse *Pax* genes and human syndromes associated with *PAX* mutations are indicated; *un*, *undulated*; *Krd*, *kidney and retinal defects*; *Sp*, *Splotch*; *Sey*, *Small eye*; KO, mouse knockout available. PAX proteins involved in human cancers are indicated.

and spatial pattern of PAX3 expression is guided by a complex array of separable regulatory elements.

1.1.3 PAX3 protein structure

Nine mammalian PAX proteins exist and, based on their sequence homology and structural characteristics, can be classified into four subgroups (Walther et al. 1991; Fig. 1-2). The defining feature of PAX transcription factors is the PD. Crystal structures of the Prd (Xu et al. 1995) and PAX6 (Xu et al. 1999) PDs in complex with DNA show that the domain consists of two helix-turn-helix (HTH) subdomains that are preceded by a β-hairpin motif and joined by a flexible linker. Alignment of PD sequences from various organisms reveals extensive conservation at the amino acid level (Balczarek et al. 1997). The PAX3 PD shares the greatest sequence identity with the PAX7 and Prd PDs, at 95% and 80%, respectively (Walther et al. 1991). Studies of Prd interaction with the *e5* sequence (5'-CACGATTAGCACCGTTCCGCTC) of the *even-skipped* (*eve*) promoter (Hoey and Levine 1988; Treisman et al. 1989) provided evidence that the Prd PD possessed DNA binding ability, and further analysis of Prd-*eve* interaction revealed that the PD behaved as an independent DNA-binding module with distinct sequence specificity (Treisman et al. 1991). Subsequent studies revealed the bipartite nature of both the PD and its target sequences and showed that the two PD subdomains make important contributions to DNA binding affinity and specificity (Czerny et al. 1993).

PAX proteins in groups III (PAX3/7) and IV (PAX4/6) feature a second DNA binding structure, the homeodomain (HD), while group II proteins (PAX2/5/8) contain a rudimentary structure that resembles the amino-terminal extension and first α-helix of the

HD (see Fig. 1-2). The PAX3 HD belongs to the Prd-type subclass defined by HDs from the *Drosophila* Prd, Gsb-proximal and Gsb-distal proteins (Bopp et al. 1986). Prd-type HDs feature invariant residues at positions 26, 27, 32, 44, 46, and 54 (Schneitz et al. 1993) and possess one of serine, glutamine, or lysine at position 50, a major determinant of HD target specificity (Hanes and Brent 1989; Treisman et al. 1989). Ser50 HDs bind preferentially to the palindromic *P2* sequence (Wilson et al. 1993), which facilitates cooperative HD dimerization; however, these HDs can also bind to sequences featuring a single recognition motif, exemplified by the *P1/2* sequence (see Table 6). The PAX3 HD and its properties will be discussed in greater detail in Chapter 5.

Separating the PD and HD of PAX3 is a flexible, 56-aa linker. During the sequencing of the HuP2 clone, Burri et al. (1989) observed that the region immediately carboxy-terminal to the PD shared significant similarity with other Prd-class proteins, including conservation of a cluster of charged residues following the PD and a HSIDGILG/S octapeptide motif. The octapeptide, which is present in group I, II, and III PAX proteins, is closely related to the Engrailed homology-1 (Eh-1) motif (Smith and Jaynes 1996). Both motifs are present in a number of other DNA-binding proteins (Copley 2005) and are characterized by a core LxxLL sequence, where L is normally leucine or isoleucine and x is any amino acid except proline. The LxxLL motif, also known as the NR box, forms a short amphipathic α-helix and is commonly regarded as an interface for protein-protein interaction, having been shown to mediate interactions between nuclear receptors and their co-activators (Savkur and Burris 2004). Also present in the interdomain linker is a stretch of negatively charged residues preceding the HD, which is another feature of Prd-class homeoproteins (Bopp et al. 1986; Schneitz et al.

1993), and a positively charged nuclear localization signal (NLS) adjacent to the HD. The carboxy-terminal region of PAX3 (and all PAX proteins) contains a Ser-Thr-Pro-rich transactivation domain (Chalepakis et al. 1994b). Additionally, some studies suggest that portions of the amino-terminal region of PAX3 (Chalepakis et al. 1994b; Bennicelli et al. 1996) and the HD (Bennicelli et al. 1996) possess transcriptional inhibitory activity, although direct repression of PAX3 target genes in the context of an endogenous system has yet to be demonstrated.

1.1.4 *Pax3* expression during development

Studies in mice revealed that *Pax3* is expressed during embryonic days 8-15, with peak expression occurring between 9-12 days post coitum (dpc) (Goulding et al. 1991). *In situ* hybridization demonstrated that *Pax3* transcripts, first observed at 8.5-9 dpc, were found in the dorsal segment of the neural tube. At least two transcription factors have been identified as potential regulators of *Pax3* expression: Mankoo et al. (1999) demonstrated that limb buds in mice lacking the MOX2 homeoprotein possessed low levels of *Pax3* expression, while Harris et al. (2002) showed that MYC family transcription factors could bind to a non-canonical E-box motif in the *Pax3* promoter and activate expression of a linked reporter gene. Additionally, transgenic analysis of the 1.6 kb upstream regulatory sequence (see Section 1.1.2) led to the identification of two neural crest enhancer elements, one of which appeared to be regulated by the TEAD2 transcription factor (Milewski et al. 2004). Despite these observations, we still have a limited understanding of how PAX3 expression is actively regulated *in vivo*.

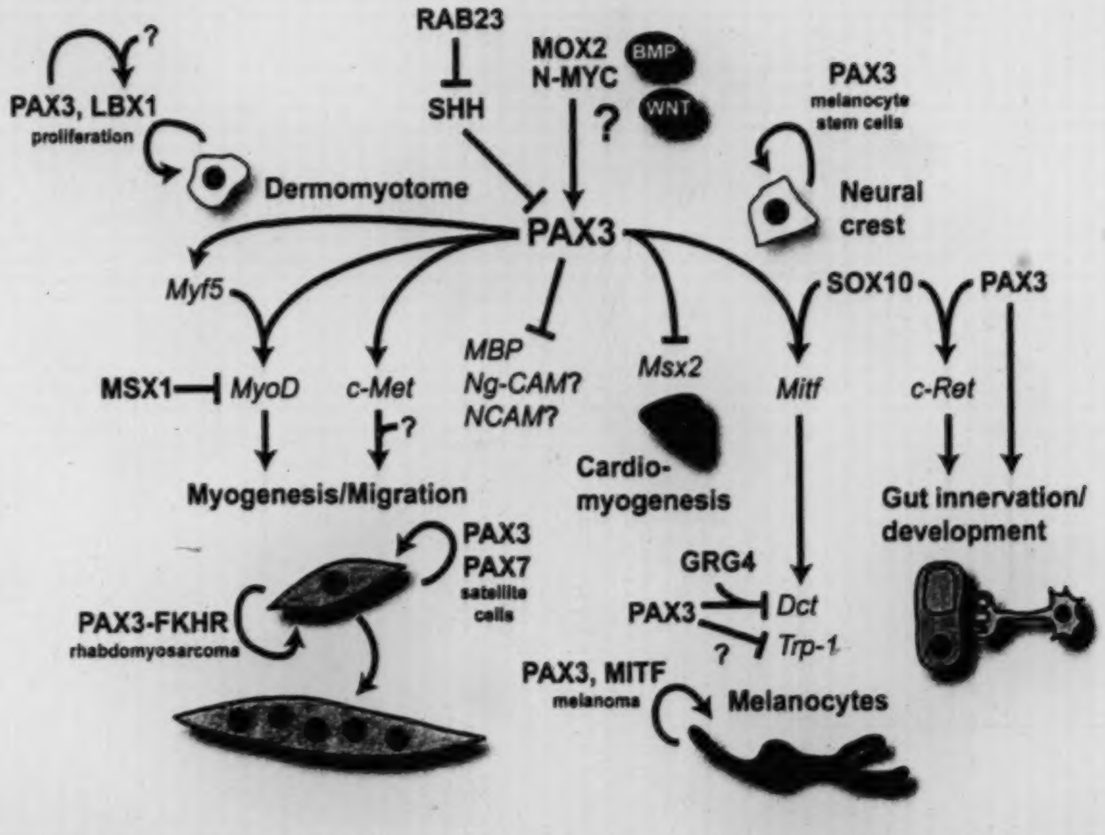

PAX3, LBX1
proliferation
?
Dermomyotome
RAB23
SHH
MOX2
N-MYC
BMP
WNT
?
PAX3
melanocyte stem cells
Neural crest
PAX3
Myf5
MSX1
MyoD
c-Met
?
MBP
Ng-CAM?
NCAM?
Msx2
Mitf
SOX10
PAX3
c-Ret
Myogenesis/Migration
Cardio-myogenesis
Gut innervation/ development
PAX3
PAX7
satellite cells
PAX3-FKHR
rhabdomyosarcoma
GRG4
PAX3
Dct
Trp-1
?
PAX3, MITF
melanoma
Melanocytes

Figure 1-3. Developmental pathways involving PAX3. PAX3 plays a pivotal role in the regulation of myogenesis, cardiomyogenesis, melanogenesis, neurogenesis, and gut development. PAX3 is responsible for activating expression of master regulators of the myogenic (*MyoD*) and melanogenic (*Mitf*) pathways and is also involved in maintaining populations of progenitor cells in these pathways (grey cells). Deregulated expression of PAX3 or derivatives contributes to human cancers, including rhabdomyosarcoma and melanoma. Putative upstream regulators of *Pax3* expression are indicated at top. Arrowheads indicate activation; flat lines indicate repression. See Sections 1.1.4 and 1.1.6 for details.

Pax3 expression in the developing spinal cord is regulated antagonistically by ventralizing and dorsalizing signals (see Fig. 1-3). Initial observations in avian embryos suggested that signals from the notochord restricted *Pax3* expression to the dorsal neural tube (Goulding et al. 1993) and later studies demonstrated a role for Sonic hedgehog (SHH), a vertebrate patterning protein secreted by the notochord, in suppressing *Pax3* expression in the ventral neural tube (Chiang et al. 1996). PAX3 itself is involved in neural tube patterning, as transgenic expression of *Pax3* throughout the neural tube early in development suppresses floor plate differentiation (Tremblay et al. 1996), although ectopic *Pax3* expression is not sufficient to cause dorsalization of ventral cells. *Pax3* also responds to dorsalizing signals from the overlying ectoderm. Bone morphogenetic proteins (BMPs) 4 and 7 were shown to enhance *Pax3* expression in the avian neural plate (Liem et al. 1995); however, later studies demonstrated that high levels of BMP4 precluded *Pax3* expression in the dermomyotome (Reshef et al. 1998). Another putative dorsalizing signal is the product of the *open brain* gene, which silences SHH activity in dorsal cells and upregulates *Pax3* expression in the developing spinal cord (Eggenschwiler and Anderson 2000; Eggenschwiler et al. 2001). In addition, Wnt signaling from posterior nonaxial mesoderm appears to restrict *Pax3* expression to the posterior neural plate (Bang et al. 1999), although basic fibroblast growth factor and retinoic acid are also capable of inducing posteriorization of *Pax3* expression (Bang et al. 1997). Together, these observations indicate that *Pax3* expression is spatiotemporally controlled through a variety of transcriptional regulatory proteins and signal transduction pathways.

1.1.5 PAX3-related diseases

Following the characterization of the *Pax* genes, mutations in the murine *Pax1* (Balling et al. 1988), *Pax3* (Epstein et al. 1991b), and *Pax6* (Hill et al. 1991; Ton et al. 1991) genes and the human *PAX3* (Baldwin et al. 1992; Morell et al. 1992; Tassabehji et al. 1992) and *PAX6* (Glaser et al. 1992) genes were found to be responsible for phenotypes associated with early developmental abnormalities. To date, each PAX protein has been shown, either in the context of a mouse knockout or a naturally occurring mutation or deletion, to play key roles in various developmental processes (Mansouri et al. 1996; Dahl et al. 1997; Underhill 2000; Chi and Epstein 2002; see Fig. 1-2). *PAX* genes are semi-dominant and loss-of-function mutations in the heterozygous state lead to haploinsufficiency (Nutt and Busslinger 1999). In addition, *PAX* genes are dosage-sensitive, exemplified by the phenotypic similarity between mice in which *PAX6* is overexpressed and *small eye* mice, which carry a heterozygous null mutation of *Pax6* (Schedl et al. 1996).

Lesions of the *PAX3* gene lead to human Waardenburg syndrome type 1 (WS1; MIM 193500). There are four clinical subtypes of WS, each distinct but with overlapping phenotypes (Read and Newton 1997). The prevalence of WS1 cases is difficult to estimate due to phenotypic variability, but is considered anywhere from 1/15,000 to 1/42,000 in the general population. WS1 and WS2 are believed to be equally prevalent, while WS3 and WS4 are rarer in occurrence. The primary clinical features of WS1 include dystopia canthorum, pigmentary anomalies, and sensorineural deafness. Dystopia canthorum is defined as medial fusion of the inner eyelids and outward displacement of the inner canthi, resulting in abnormally increased inner and outer canthal distances. Dystopia is seen in approximately 99% of WS1 patients and thus is the main diagnostic feature of WS1. Other facial characteristics often observed in WS1 patients include a broad nasal root, synophrys, and hypoplastic alae nasi. Pigmentary defects in WS1 normally manifest in hair, eyes, and skin; a white forelock, which can be present at birth or appear later, is a common feature. Ocular pigmentary anomalies include iris heterochromia, which can be complete (each eye is a different color) or segmental (one iris shows two different colors). Bright blue irises ("Waardenburg blue eyes") are common in WS1 and WS2 patients, although seemingly more prevalent in WS1 individuals. Sensorineural deafness associated with WS1 is congenital and nonprogressive and the extent of hearing loss can be variable. Deafness can be bilateral or unilateral and is believed to result from the failure of melanocytes to populate the inner ear. Sprengel's shoulder, cleft lip or palate, upper limb contractures (associated with WS3), and in severe cases, neural tube defects (NTDs), have also been noted in WS1 patients. Phenotypic variability among WS1 individuals is common, even within families, suggesting the existence of modifier loci or gene-environment interactions.

Since the discovery that *PAX3* was genetically linked to WS1 (Farrer et al. 1994), numerous studies have provided evidence supporting that disruption of *PAX3* causes WS1. At the time of writing, sixty-five unique disease-associated lesions of *PAX3*, including missense and nonsense mutations, small and whole-gene deletions, small insertions, and splice site aberrations, have been characterized, according to the Human Gene Mutation Database (www.hgmd.cf.ac.uk) (Fig. 1-4). The majority of these lesions lead to WS1 and so far *PAX3* is the only gene known to be associated with the disease. Known missense mutations map to regions of *PAX3* that code for the DNA-binding

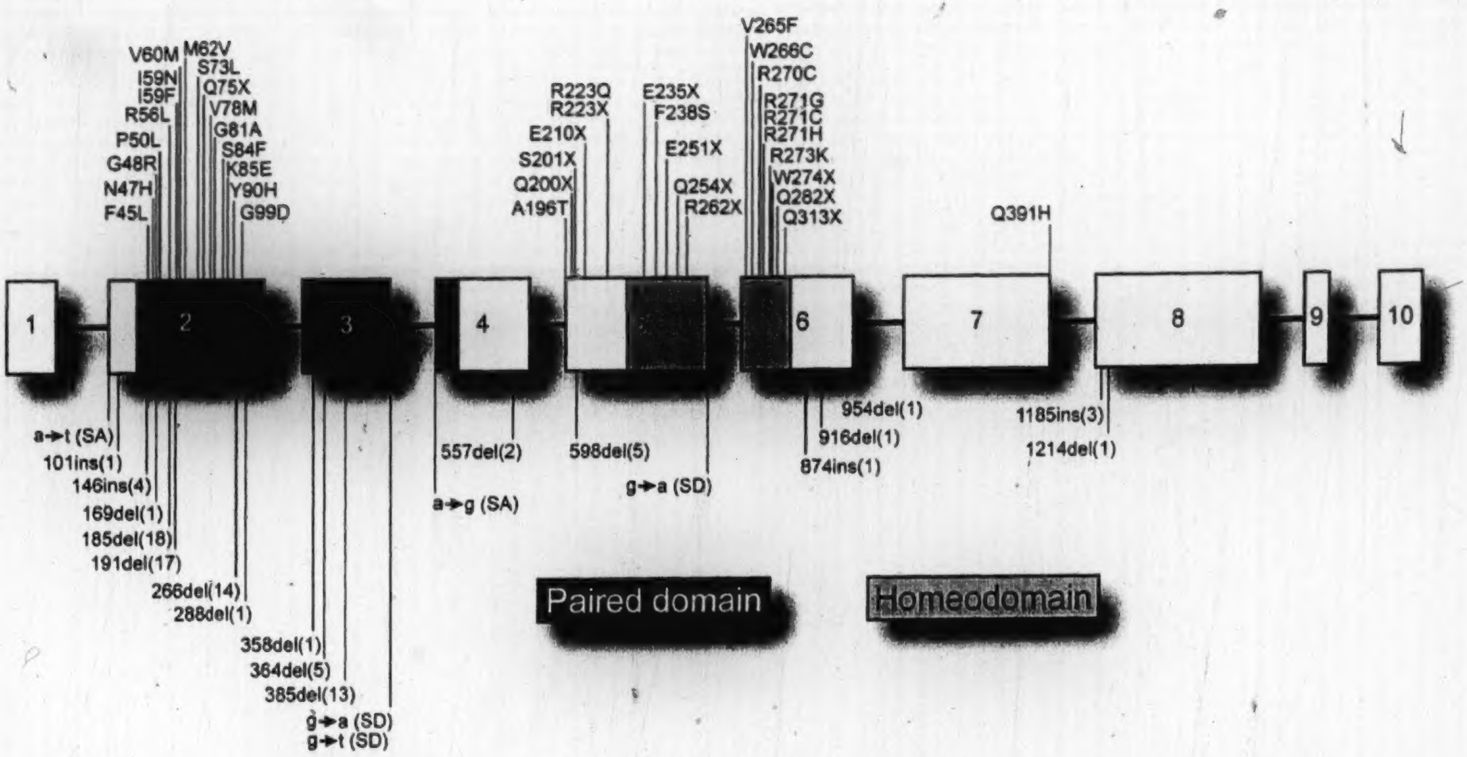

V60M
M62V
I59N
S73L
I59F
Q75X
R56L
V78M
G81A
P50L
S84F
G48R
K85E
N47H
Y90H
F45L
G99D
R223Q
R223X
E235X
F238S
E210X
E251X
S201X
Q200X
Q254X
A196T
R262X
V265F
W266C
R270C
R271G
R271C
R271H
R273K
W274X
Q282X
Q313X
Q391H
1
2
3
4
6
7
8
9
10
a→t (SA)
101ins(1)
146ins(4)
169del(1)
185del(18)
191del(17)
266del(14)
288del(1)
358del(1)
364del(5)
385del(13)
g→a (SD)
g→t (SD)
557del(2)
a→g (SA)
598del(5)
g→a (SD)
954del(1)
916del(1)
874ins(1)
1185ins(3)
1214del(1)
Paired domain
Homeodomain

Figure 1-4. Positions of *PAX3* lesions associated with Waardenburg syndrome. Exons are numbered and locations of the paired domain and homeodomain are shown. Nonsense (indicated by 'X') and missense mutations are listed at the top; small deletions, insertions, and splice acceptor (SA) and splice donor (SD) mutations are shown at the bottom.

domains, while the remainder of *PAX3* mutations and deletions/insertions occur throughout exons 2-6. Curiously, very few disease-causing mutations have been found in the 3' region of *PAX3*, which encodes the transcription activation domain. This suggests that either single amino acid changes are tolerated in the domain or that mutations in this part of the gene contribute to embryonic lethality and thus would not be identified in current mutation screens. Other syndromes caused by PAX3 defects include WS3 (MIM 148820) and craniofacial-deafness-hand syndrome (CDHS) (MIM 122880).

Mutations of *Pax3* cause the murine *Splotch* (*Sp*) phenotype and provide an animal model for studying human *PAX3*-related disorders. The *Sp* and WS1 phenotypes are analogous, although differences do exist. Most notably, *Sp* mice do not exhibit hearing loss (Steel and Smith 1992) or craniofacial defects, although the latter is influenced by genetic background (Asher et al. 1996a). The original *Sp* allele results from the mutation of a splice acceptor site at the end of intron 3, leading to a protein truncated after the PD (Epstein et al. 1993). The Sp^{2H} allele features a 32-bp deletion within the HD (Epstein et al. 1991b), while the *Splotch-retarded* and Sp^{4H} alleles are characterized by whole-gene deletions and result in persistent growth retardation in heterozygotes (Epstein et al. 1991a). A Gly→Arg mutation in the PD β-hairpin causes the *Splotch-delayed* (Sp^{d}) phenotype, which is the mildest of the *Sp* alleles in the heterozygous state (Vogan et al. 1993). Homozygous *Sp* mice are not viable and exhibit NTDs, a decrease in the size and number of spinal ganglia, and, in the case of the *Sp* and Sp^{2H} alleles, persistent truncus arteriosis and reduction of Schwann cell density.

PAX3 is expressed in a variety of cancers, including rhabdomyosarcoma, melanoma, and Ewing's sarcoma (Bernasconi et al. 1996; Schulte et al. 1997; Frascella et al. 1998; Barr et al. 1999; Scholl et al. 2001). Chromosomal translocations involving group II and III *PAX* genes create potent transcriptional activators involved in several human neoplasms (Barr 1997; Robson et al. 2006). A balanced translocation involving the *PAX3* and *FKHR* genes is a frequent feature of alveolar rhabdomyosarcoma, a rare childhood cancer. In these individuals, the t(2;13)(q35;q14) translocation fuses the amino-terminal region of PAX3, including the DNA-binding domains, to the carboxy-terminal region of FOXO1A, which possesses a strong transactivation domain (Galili et al. 1993; Shapiro et al. 1993). The resulting fusion protein (PAX3-FKHR) is characterized by gain-of-function and oncogenic properties. Recent studies have demonstrated that PAX3-FKHR can rescue PAX3 mutant phenotypes, although abnormal myogenic activity was also observed (Relaix et al. 2003; Keller et al. 2004a; Keller et al. 2004b). Finally, a translocation involving PAX3 and the nuclear receptor co-activator NCOA1 also causes rhabdomyosarcoma and provides further evidence that the PAX3 moiety is required for pathogenesis (Wachtel et al. 2004).

1.1.6 Role of PAX3 in developmental pathways

The findings that *Sp* mice lack limb musculature (Franz et al. 1993) and do not express PAX3 in migratory cell populations or the developing limb (Bober et al. 1994) pointed to a role for PAX3 in the myogenic pathway. Support was provided by studies demonstrating that ectopically expressed PAX3 induced expression of myogenic markers, including *Myf5* and *MyoD*, in neuroepithelium and surface ectoderm (Maroto et al. 1997). Compound $Pax3^{-/-}/Myf5^{-/-}$ mice showed a more severe phenotype than null mutants of either gene alone, including an absence of limb and body musculature (Tajbakhsh et al.

1997), suggesting that PAX3 and MYF5 operate upstream of *MyoD*, the so-called 'master regulator' of myogenesis. It was later shown that PAX3 directly regulates *Myf5* expression in the hypaxial somite and derivative muscle tissue (Bajard et al. 2006), placing PAX3 at the top of the myogenic hierarchy. Mice harboring null mutations of *Mox2*, which encodes a suspected *Pax3* regulator (see Section 1.1.4), displayed limb muscle defects concomitant with downregulation of *Pax3*, *Myf5*, and *c-Met*, but with normal levels of *MyoD* expression (Mankoo et al. 1999). These observations suggest that regulation of *MyoD* by PAX3 operates through mechanisms that differ between limb and axial muscle precursors. Studies of *Sp* mice revealed an absence of cells expressing the c-MET receptor which, among other functions, promotes the migration of epithelial and endothelial cells during embryonic development (Daston et al. 1996; Yang et al. 1996). These results, together with those suggesting that PAX3 regulates expression of *c-Met* (Epstein et al. 1996), help account for the failure of muscle precursor cells to populate the limb buds of *Sp* mice. The importance of PAX3 in limb muscle development was emphasized upon substitution of *Pax3* for *Pax7* in mice, which caused deficiencies in limb muscle precursor migration despite normal neural tube and somite development (Relaix et al. 2004). Another myogenic pathway involving PAX3 parallels a regulatory loop controlling eye development in *Drosophila*, in which PAX3 synergistically cooperates with the DACH2, EYA2, and SIX1 proteins to regulate muscle development (Heanue et al. 1999). In addition, PAX3, together with the LBX1 homeoprotein, helps control the proliferation and expansion of myoblast populations (Mennerich and Braun 2001). PAX3 also appears to play a role in cardiomyogenesis, as *Sp* mice exhibit defects in cardiac neural crest migration and aorticopulmonary septation (Kwang et al. 2002). In addition, studies using a transgenic construct containing the *Pax3* proximal promoter showed that *Pax3* is expressed in cells that contribute to cardiac structures, including the aortic arch, proximal head vessels, and pulmonary artery (Li et al. 2000). Collectively, these observations demonstrate that PAX3 is an integral factor in the control of myogenic processes.

Pax3 is expressed in a population of neural crest-derived cells, some of which are destined to differentiate into melanoblasts. *In vitro* studies have shown that PAX3 can bind to a recognition element in the *microphthalmia-associated transcription factor* (*MITF*) promoter and activate expression of a promoter-linked reporter construct (Watanabe et al. 1998; Corry and Underhill 2005a). MITF is responsible for activating expression of the *tyrosinase-related protein-1* (*Trp-1*) and *dopachrome tautomerase* (*Dct*; also known as *Trp-2*) genes, driving melanocyte differentiation (Goding 2000). *Sp* mice were shown to exhibit normal migration but diminished numbers of melanoblast precursor cells, while *Mitf* mutant mice display a reduction in committed melanoblasts (Hornyak et al. 2001). These results suggest a pathway in which PAX3 is responsible for the expansion of a population of melanoblast progenitor cells, while MITF ensures the survival and rate of differentiation of committed melanoblasts. Later studies demonstrated that PAX3 could directly repress the *Dct* gene (Lang et al. 2005), invoking a model in which PAX3, at the top of the melanogenic pathway, functions to specify melanocyte commitment while simultaneously inhibiting terminal differentiation. Thus, PAX3 appears to act as a genetic switch that commits cell populations to a specific fate by activating 'master regulators,' such as *Myf5*/*MyoD* and *Mitf*, in the myogenic and melanogenic hierarchies, respectively (see Fig. 1-3 for a summary).

PAX3 also functions postnatally in several pathways. In differentiated primary cultures, *Pax3* is expressed in granule neurons, Purkinje neurons, and astroglia (Kioussi and Gruss 1994). In addition, the use of genetic fate mapping techniques suggests PAX3 plays a role in the development of hindgut colonic epithelium and the genitourinary tract (Engleka et al. 2005). PAX3 also functions in melanocyte stem cells where it is thought to maintain an undifferentiated state (Lang et al. 2005). Lastly, a requirement for PAX3 in a subset of satellite cells (Kuang et al. 2006; Relaix et al. 2006) points to a role in adult myogenesis.

1.1.7 PAX3 target genes

Goulding et al. (1991) demonstrated that PAX3 could bind to the *e5* sequence, which features both PD (5'-GTTCC) and HD (5'-ATTA) recognition motifs and is bound by the Prd PD and HD (Hoey and Levine 1988; Treisman et al. 1989; see Section 1.1.3). Importantly, mutation of the *e5* GTTCC motif abrogated PAX3 binding, while mutation of the HD motif severely reduced PAX3 affinity for the sequence. These results implied that efficient recognition of *e5* by PAX3 requires participation of both the PD and the HD. Subsequent independent analyses of PAX3 PD DNA binding specificity identified a sequence, termed *Nf3'*, to which PAX3 bound with high affinity (Chalepakis and Gruss 1995; Epstein et al. 1995). However, *Nf3'* differs from *e5* in the composition of the PD site (5'-GTCAC vs. 5'-GTTCC) and does not contain a HD-binding motif.

Based on these findings and supporting genetic evidence, Epstein et al. (1996) identified two motifs, MET1 and MET2, in the promoter of the *c-MET* gene, which encodes a tyrosine kinase receptor for the hepatocyte growth factor/scatter factor

(HGF/SF) (Bottaro et al. 1991) and is a marker of migratory neural crest cells (see Section 1.1.6). Although it features a change at a highly conserved position within the core PD-binding site, the MET1 motif was bound by the PAX3 PD with high affinity (Epstein et al. 1996); however, the MET2 site, which varies at the same position as MET1, was bound by the PD with low affinity. The Epstein group also provided evidence that PAX3 controls the expression of the *c-RET* gene (Lang et al. 2000), which also encodes a tyrosine kinase receptor and plays a role in the development of neural crest-derived enteric ganglia precursors (Pachnis et al. 1993; Romeo et al. 1994). Conserved recognition elements for PAX3 and SOX10, a high-mobility-group (HMG) family transcription factor, were discovered in the *c-RET* promoter (Lang et al. 2000) and both factors were shown to bind to the promoter and synergistically activate expression of a linked reporter gene. The same group later demonstrated that PAX3 and SOX10 interact through the PD and HMG box, respectively, and that this interaction underlies the synergistic effects on *c-RET* expression (Lang and Epstein 2003).

PAX3 regulates melanogenesis by activating expression of *MITF* and repressing expression of *Dct* (Lang et al. 2005; see Section 1.1.6). PAX3 binds to an element in the *MITF* promoter and activates transcription of a reporter gene linked to a portion of the promoter containing this sequence (Watanabe et al. 1998). The element contains putative binding motifs for both the PD and HD, and it was subsequently shown that the PD motif and a non-canonical HD-recognition site were necessary for high affinity binding to the *MITF* promoter element (Corry and Underhill 2005a). As with *c-RET*, PAX3 and SOX10 have been shown to synergistically activate *MITF* expression (Bondurand et al. 2000; Potterf et al. 2000). However, these and other studies demonstrated that SOX10 can activate expression of a *MITF* promoter-linked construct in the absence of PAX3 (see also Lee et al. 2000; Verastegui et al. 2000), while other data suggest that PAX3 and SOX10 do not synergistically activate *MITF* expression (Lee et al. 2000; Verastegui et al. 2000). Moreover, the *MITF* and *c-RET* promoters differ substantially in binding site composition – *MITF* possesses multiple putative binding motifs for PAX3 and SOX10 (Bondurand et al. 2000; Lee et al. 2000; Potterf et al. 2000), while *c-RET* features single recognition elements for each protein (Lang et al. 2000; Lang and Epstein 2003). Thus, differences in promoter structure may dictate if and how PAX3 and SOX10 interact, and whether synergistic activation is achieved.

In mice, PAX3 was shown to repress the expression of the *Dct* gene by competing with MITF for an M-box motif in the *Dct* promoter and repressing expression of a *Dct* promoter-linked reporter construct in a dose-dependent manner (Lang et al. 2005). Although PAX3 bound with low affinity to the *Dct* regulatory region, chromatin immunoprecipitation (ChIP) studies demonstrated that PAX3 is able to interact with the region *in vivo* and that PAX3 binding precludes MITF-directed expression of *Dct*. However, it is unlikely that PAX3 is the sole factor required to repress *Dct* expression, as mutation of a putative TCF/LEF site in the *Dct* enhancer abrogated PAX3-mediated repression, possibly by interfering with the formation of a complex that includes the GRG4 transcriptional co-repressor (see below). PAX3 is also thought to regulate expression of the *Trp-1* gene (see Section 1.1.6). A melanocyte specific factor (MSF), later revealed to be PAX3 (Galibert et al. 1999), was shown to bind to two motifs in the *Trp-1* promoter (Yavuzer and Goding 1994). These melanocyte-specific initiator elements, termed MSEu and MSEi, each feature a 5'-GTGTGA core sequence that resembles a noncanonical PAX3-recognition element on the complementary strand. PAX3 was shown to interact with each motif in DNA binding assays and efficiently transactivated expression of a *Trp-1* promoter-linked reporter construct (Yavuzer and Goding 1994; Galibert et al. 1999). However, the minimal region of the *Trp-1* promoter required for expression in B16F10 mouse melanoma cells includes only the MSEi motif (Lowings et al. 1992), while the MSEu motif has also been shown to participate in downregulation of *Trp-1* expression (Yavuzer and Goding 1994). Thus, it remains unclear how, if at all, PAX3 is involved in regulating *Trp-1* transcription. Furthermore, given the role of *Trp-1* (and *Dct*) in promoting differentiation of melanocytes, it seems counterintuitive that PAX3 would activate *Trp-1* and repress *Dct* expression, unless other factors not present in the *Trp-1* studies, such as GRG4 and TCF/LEF, are required for its repression.

During myogenesis, PAX3 activates expression of the *MYF5* gene (Bajard et al. 2006; see Section 1.1.6). Studies using ChIP showed that PAX3 interacts with a consensus recognition element in the *MYF5* promoter, while expression of a transgene linked to a segment of the *MYF5* upstream regulatory region containing this element was directed to the mature somite and limb buds, providing the first *in vivo* evidence of a gene directly regulated by PAX3. Based on these studies, however, it was apparent that PAX3 is not the sole regulator of *MYF5* expression throughout muscle development and, indeed, it was recently demonstrated that members of the SIX homeoprotein family coordinately regulate *MYF5* expression with PAX3 (Giordani et al. 2007). Similar to the *MITF* promoter, the PAX3-recognition element of *MYF5* contains a composite sequence featuring binding motifs for a PD and HD (Bajard et al. 2006) and, like *MITF*, it was shown that both motifs are required for high affinity binding to the *MYF5* site (Buchberger et al. 2007). PAX3 also has a repressive function in myogenesis, as Kwang et al. (2002) showed that defects in cardiac neural crest migration in *Sp* mice likely result from the failure of PAX3 to repress expression of the MSX2 homeoprotein. A conserved 5'-GTCAC motif lies in the *Msx2* promoter and mutation of this motif abrogated PAX3 binding and caused expanded expression of an *Msx2-LacZ* transgene *in vivo*.

Microarray analysis using RNA from PAX3-overexpressing cell lines implicated *TGFβ2*, an important gene in neural tube development (Sanford et al. 1997), as another potential target of PAX3 (Mayanil et al. 2001). Through ChIP and reporter gene assays, this group demonstrated that PAX3 binds to the *TGFβ2* promoter and activates its expression. However, the mechanistic basis of *TGFβ2* regulation by PAX3 is unclear, as multiple putative PD and HD binding motifs were shown to be distributed across over half a kb of the *TGFβ2* upstream regulatory region. Furthermore, both activating and repressive elements were found in the *TGFβ2* promoter, suggesting that PAX3 may require the participation of additional factors to properly regulate this gene. PAX3 has been shown to regulate other genes involved in development of the nervous system. Kioussi et al. (1995) demonstrated repression of a reporter gene linked to the *myelin basic protein* (*MBP*) promoter by PAX3, suggesting a role for PAX3 in Schwann cell development. PAX3 was also shown to repress basal transcription of a reporter gene fused to the *neural cell adhesion molecule* (*N-CAM*) promoter (Chalepakis et al. 1994b); however, no binding of PAX3 to an oligonucleotide containing the *N-CAM* promoter sequence was observed. PAX3 also binds to a silencer element in the *neuron-glia cell adhesion molecule* (*Ng-CAM*) promoter and repressed a reporter gene linked to the promoter (Kallunki et al. 1995). These observations suggest that PAX3 plays a predominantly repressive role in neuronal development, although additional studies are required to further define its role in this process.

The PAX3-FKHR fusion protein is a more potent transactivator than PAX3 (Fredericks et al. 1995), but since it features intact PAX3 DNA-binding domains, one might expect PAX3-FKHR to interact with the same targets as PAX3. However, this does not appear to be the case, as PAX3 overexpression does not produce an oncogenic phenotype (Tremblay et al. 1996; Wada et al. 1997), implying that the two proteins regulate distinct sets of genes. Indeed, it was shown that PAX3-FKHR can activate expression of the *platelet-derived growth factor-alpha receptor* (Epstein et al. 1998) and *myogenin* (Zhang and Wang 2007) genes, which are not considered direct targets of PAX3. Furthermore, microarray analyses have identified genes and developmental pathways that are differentially regulated by PAX3 and PAX3-FKHR (Khan et al. 1999; Begum et al. 2005). Mice expressing knock-in PAX3-FKHR alleles display a number of developmental abnormalities (Lagutina et al. 2002; Keller et al. 2004b), but tumor formation was only observed at high frequency upon disruption of the Ink4a/ARF and Trp53 pathways (Keller et al. 2004a), suggesting that, although PAX3-FKHR plays a role in the onset of alveolar rhabdomyosarcoma, other factors are required for a malignant phenotype. Finally, in contrast to PAX3, PAX3-FKHR is unresponsive to the repressive effects of the DAXX transcriptional co-repressor (Hollenbach et al. 1999; see Section 1.1.8), suggesting that the fusion protein is capable of subjugating the mechanisms that control PAX3-mediated gene regulation.

1.1.8 PAX3 protein-protein interactions

PAX3 interacts with a number of different proteins that are thought to modulate its transcriptional activity, either by enhancement or repression of target gene expression. Discussed above is the interaction between PAX3 and the SOX10 transcription factor that is thought to synergistically enhance activation of the *c-RET* and *MITF* genes (see Section 1.1.7). Since PAX3 and SOX10 are also involved in regulation of glial cell development (Kioussi et al. 1995; Kuhlbrodt et al. 1998), these results collectively support a role for both factors in the cooperative regulation of genes in tissues characterized by *Sox10* and *Pax3* expression. In these cases, however, it remains unclear if the SOX partner plays a role in target specificity beyond an alteration to local DNA or chromatin architecture, which is one function of SOX proteins (Pevny and Lovell-Badge 1997). The POU domain transcription factor BRN2 was also demonstrated to interact with PAX3 and enhance activation of a GAL4-responsive reporter gene by PAX3-GAL4 (Smit et al. 2000). The involvement of BRN2 and PAX3 in melanogenesis and melanoma (Eisen et al. 1995; Scholl et al. 2001) supports a possible role for a functional interaction between the two proteins. The MOX2 transcription factor, which is thought to activate *Pax3* expression (see Section 1.1.4), is another putative PAX3-interacting protein. Stamataki et al. (2001) demonstrated that the two proteins physically interact *in vitro* through their respective HDs; however, no *in vivo* interaction was detected and, to date, no data regarding the effects of MOX2 on PAX3-mediated gene expression have been reported. Together, these results indicate that, although PAX3 likely regulates gene expression by interacting with other transcription factors at promoter sequences, further studies are needed to clarify the nature of these interactions and their relevance in physiological systems *in vivo*.

PAX3 also interacts with proteins that exert repressive effects on transcription. The HIR histone cell cycle regulation defective homolog A (HIRA) protein, which functions as a histone chaperone and is involved in cellular senescence (Lorain et al. 1998; Zhang et al. 2005), was shown to interact with a region of PAX3 containing the HD (Magnaghi et al. 1998). Retinoblastoma (pRB) family members, including pRB, p107, and p130, were also shown to interact with the HD of PAX3 and related Prd-class homeoproteins (Wiggan et al. 1998). Similarly, the PAX3 HD, together with the octapeptide and adjacent regions, was shown to mediate PAX3 interaction with DAXX, a death-associated protein that can also act as transcriptional co-repressor (Hollenbach et al. 1999). The KRAB-associated protein 1 (KAP1) and heterochromatin protein 1 (HP1) were shown to competitively interact with PAX3, regulating its repressive activity in the presence of an artificial reporter construct (Hsieh et al. 2006). The GRG4 protein, a murine homolog of the Groucho (Gro)/transducin-like enhancer of split (TLE) transcriptional co-repressors, also interacts with PAX3 and is likely responsible for enhancing repression of the *Dct* gene (Lang et al. 2005) (see above). PAX3 interaction with GRG4 appears to require the presence of LEF1, as mutation of the TCF/LEF binding site in the *Dct* promoter or ectopic expression of β-catenin, which complexes with LEF1, abrogated *Dct* repression by PAX3. Importantly, the unifying theme of these interactions is that HIRA, DAXX, pRB family proteins, KAP1, HP1, and TLE/GRG proteins all function in transcriptional downregulation or silencing. Lastly, the calcium-binding protein calmyrin was shown to interact with the amino-terminal region of PAX3, including the PD and octapeptide, and repressed transactivation of a PAX3-responsive reporter construct (Hollenbach et al. 2002). Together, these observations suggest that PAX3, in addition to its function as a transcriptional activator, can mediate gene repression upon interaction with a variety of proteins. However, since none of these interactions have been demonstrated in the context of the endogenous proteins, a more detailed analysis of their underlying characteristics needs to be performed.

1.1.9 Functional interactions involving the paired domain and homeodomain

Past studies of PD and HD DNA binding behavior suggests that, upon cooperative binding to a single DNA template, the two domains are situated in close proximity to each other. The *e5* sequence (see Section 1.1.3) features putative PD- and HD-binding motifs separated by 5 bp and was shown in footprinting experiments to be bound by Prd. *In vitro* examinations of Prd binding to the *PH0* sequence (Jun and Desplan 1996), which features juxtaposed PD- and HD-binding motifs, suggested that the domains would be arranged such that the second α-helix of the amino-terminal PD subdomain would contact the amino-terminal region of the HD. At the same time, it was shown that the presence of both the PD and HD were necessary for rescue of *prd* mutants *in vivo* (Bertuccioli et al. 1996; Miskiewicz et al. 1996). Similarly, binding of both domains to the *eve* promoter *Paired Target Element* (*PTE*), which features adjacent PD and HD recognition motifs, was shown to confer optimal expression of *eve* by Prd (Fujioka et al. 1996) and later studies demonstrated a strict requirement for spacing of the PD and HD sites within *PTE* (Lan et al. 1998). Together, these observations suggest that the PD and HD are both required for optimal binding to sequences that feature motifs for both domains and that *in* *vivo* regulation of gene expression by PD factors is dependent on the function of both domains.

Like Prd, PAX proteins also display functional interdependence between the PD and HD. Studies of human disease alleles showed that mutations in the PD and HD of PAX3 (Baldwin et al. 1995) or PAX6 (Glaser et al. 1992) produce similar phenotypes. Subsequent studies showed that the Sp^d mutation (see Section 1.1.5) affected PAX3 binding to both canonical PD- and HD-recognition motifs (Underhill et al. 1995; Underhill and Gros 1997), while a mutation in the PAX3 HD recognition helix exerts the same effect (Fortin et al. 1997). Scanning mutagenesis of PAX3 also showed that interference with one domain's function disrupted the activity of the other (Apuzzo and Gros 2002; Apuzzo et al. 2004). Together, these results suggest that the PD and HD of PAX3 are functionally interdependent and again emphasize that the integrity of both domains is required for its normal function. These results also imply that the PAX3 PD and HD physically contact each other when bound to DNA, a model supported by the proposed contact between the Prd PD and HD on *PH0* (see above) and functional interactions involving the PAX6 PD and HD (Mikkola et al. 2001; Bruun et al. 2005). Recent analysis of the putative contact region between the PAX3 PD and HD demonstrated that residues in α-helices 1 and 2 and the loop between α-helices 2 and 3 of the amino-terminal PD subdomain lie proximal to residues of the amino-terminal extension of the HD (Apuzzo and Gros 2006; Apuzzo and Gros 2007), suggesting that functional interaction between the PD and HD influences PAX3 DNA binding behavior.

The above observations suggest that intramolecular interactions involving the PD and HD of PAX proteins support functional cooperativity; however, PAX proteins also cooperate with other factors to enhance DNA binding. So far, the standard for analyzing PAX protein cooperativity has been the PAX5-ETS model, in which ETS family transcription factors are recruited to the B-cell specific *CD79a* promoter by PAX5 (Fitzsimmons et al. 1996). The PAX5 PD plays the dominant role in cooperativity and mediates high affinity binding of ETS proteins to a suboptimal recognition motif through a DNA-dependent conformational rearrangement of the ETS1-DNA complex (Garvie et al. 2001). Interestingly, optimizing the ETS site causes a reduction in PAX5 binding (Wheat et al. 1999), suggesting that the suboptimal recognition motif supports PAX5-ETS cooperativity. Interactions between PAX proteins, including PAX3 (see above), and SOX family transcription factors have also been shown to enhance cooperativity in the presence of suboptimal binding sequences (Kuhlbrodt et al. 1998; Bondurand et al. 2000; Lang et al. 2000; Potterf et al. 2000; Kamachi et al. 2001). Thus, cooperative interactions between PAX3 and transcription factors from other families may act to fine-tune PAX3-mediated activation of downstream target genes. However, the precise nature of these interactions, including the involved domain(s) and the dependence on a DNA template, requires further elucidation.

1.2 PAX3: from the test tube to the nucleus

The PAX family of transcription factors has been the focus of intense study over the past two decades and much research has centered on characterizing the role of PAX3 in the regulation of developmental processes and how mutant PAX3 alleles cause disease phenotypes (discussed above). The majority of our knowledge regarding PAX3 functional behavior, including DNA binding properties, PAX3-protein interaction, and

D

trans-regulation of target genes, has been realized from *in vitro* assays. With rapid improvements in the ability to examine the behavior of nuclear components, both in fixed and live cells, researchers are gaining a better understanding of transcription factor behavior in the context of the nucleus. Through these analyses, we are becoming increasingly aware of how intranuclear compartmentalization and mobility influence a given transcription factor's behavior and how disease-associated mutations affect these proteins *in vivo*. The following section provides an introduction to the structure and functional organization of the eukaryotic nucleus with a focus on the behavior and properties of sequence-specific transcription factors in this context. Together with the above introduction to PAX3, this section is intended to provide a suitable foundation for understanding the determinants of PAX3 behavior in the nucleus and appreciating how disease mutations contribute to PAX3 dysfunction *in vivo*.

1.3 Intranuclear organization of transcription

1.3.1 The eukaryotic nucleus

The eukaryotic nucleus is a highly organized and dynamic structure containing the cell's genetic information as well as the components required for DNA replication and repair, transcription, and post-transcriptional RNA processing (for reviews, see de Jong et al. 1996; Lamond and Earnshaw 1998; Spector 2003; Misteli 2007). *In vitro* studies have revealed a wealth of data regarding the intricacy of these processes, but the functional and spatial organization of the nucleus and its role in coordinating these processes *in vivo* remains poorly understood. Indeed, two models of nuclear organization have been speculated: a deterministic model, in which nuclear structures act as stable architectural units that regulate nuclear processes, and a self-organizing model, in which the functional architecture of the nucleus is dynamic and requires no physical scaffold per se (Misteli 2007). Both models are supported by existing evidence. The presence of filamentous structures in the nucleus, such as lamins, actin, and the hypothetical nuclear matrix, as well as relatively immobile chromosome territories, suggest that nuclear processes might be supported by a stable framework of nuclear components; however, little experimental evidence exists to endorse such a model (see Misteli 2007). Conversely, the self-organizing model is supported by the highly dynamic nature of nuclear constituents and observations demonstrating the functional and structural interdependence of nuclear bodies. For example, inhibition of mRNA synthesis causes RNA polymerase II (RNAPII) and splicing machinery to accumulate in large nuclear foci and, upon release from transcriptional inhibition, the proteins redistribute to their typical nuclear patterns (Bregman et al. 1995).

In a mammalian cell nucleus, approximately two meters of DNA must be compacted into a space about ten microns in diameter. DNA compaction is achieved through association with histone proteins, forming nucleosomes, the basic units of chromatin. Approximately 146 bp of DNA is wound helically around an octamer consisting of two copies of each histone protein, H2A, H2B, H3, and H4, forming the nucleosome core particle (Luger et al. 1997). The globular domains of the histones form the inner core of the nucleosome particle, while flexible amino-terminal tails protrude from between the helical gyres of the DNA. The interaction of linker histone H1 with internucleosomal linker DNA facilitates higher-order compaction and completes the nucleosome structure (Thomas 1999). Nucleosome arrays are highly dynamic and have been shown to be compacted into higher-order structures, although whether some of these structures are formed *in vivo* remains debatable (Hansen 2002).

Chromatin structure and associated modifications play a major role in regulating gene expression, DNA repair, and replication (see Groth et al. 2007; Li et al. 2007 for recent reviews). Two forms of higher-order chromatin exist: gene-rich euchromatin, which is loosely associated and competent for transcriptional activity, and heterochromatin, which is condensed, gene-poor, and generally inaccessible to transcriptional machinery. Within the past decade, a considerable amount of data has been collected describing the post-translational modifications of nucleosomal histone proteins, including acetylation, methylation, phosphorylation, ubiquitylation, SUMOylation, and ADP-ribosylation, and their effects on nuclear processes, including transcription and DNA repair (Berger 2002; Iizuka and Smith 2003; Peterson and Laniel 2004; van Attikum and Gasser 2005). Histone modifications are generally thought to modulate local chromatin structure by influencing internucleosomal interactions and association with chromatin-binding proteins. However, these modifications may also provide an epigenetic 'code' that is recognized and 'read' by sets of proteins or protein complexes responsible for maintaining a specific cellular state (Strahl and Allis 2000; Jenuwein and Allis 2001). Indeed, researchers have discovered an array of protein domains and modules responsible for depositing and recognizing specific histone modifications associated with alteration or maintenance of a particular genetic program (Marmorstein 2001). In addition, variants of the core and linker histones can be incorporated into nucleosomes, providing a further level of complexity in epigenetic regulation (Kamakaka and Biggins 2005; Sarma and Reinberg 2005). Epigenetic regulation of gene expression is not limited to histone modifications, however. Both DNA methylation (Weber and Schubeler 2007) and the involvement of non-coding RNA molecules in determining chromatin states (Bernstein and Allis 2005) provide additional tiers to the epigenetic regulation of transcription in eukaryotes. Together, these phenomena create a stable regulatory system that acts in conjunction with elements at the DNA sequence level to precisely control gene expression at multiple stages throughout the life of a cell.

It is a well accepted fact that the nucleus is highly organized, but the mechanistic basis of this organization remains a contentious issue. Current models suggest that individual chromosomes occupy distinct territories in the interphase nucleus, and that the nuclear space devoid of any chromatin material, termed the interchromatin compartment, forms a second functional nuclear domain (Cremer and Cremer 2001; Cremer et al. 2006). In this model, chromosome territories are functionally arranged so that coding sequences on chromatin fibers, whether active or inactive, are positioned at the territory's periphery, while non-coding sequences form the territory's interior (Kurz et al. 1996). Indeed, transcriptional activity has been found at the surface of chromosome territories (Cmarko et al. 1999; Volpi et al. 2000; Mahy et al. 2002a; Williams et al. 2002; Chambeyron and Bickmore 2004), although evidence of transcriptional activity in the chromosome territory interior also exists (Visser et al. 1998; Verschure et al. 1999; Tajbakhsh et al. 2000; Mahy et al. 2002b). In addition, the structure of chromosome territories has been demonstrated to be dependent on gene density and permissibility of the region to transcription, rather than activity of individual genes (Mahy et al. 2002a), while gene density is also known to play a role in positioning individual chromosomes within the nucleus (Croft et al. 1999).

The nuclear space left unoccupied by chromosome territories forms the interchromatin compartment, a DNA-depleted network that accommodates factors involved in processes such as DNA replication, repair, and transcription (Zirbel et al. 1993; Cremer et al. 2000). This space also contains a multitude of functionally separable domains, including splicing speckles, Cajal bodies, promyelocytic leukemia (PML) bodies, and gems, to name a few (Dundr and Misteli 2001; Spector 2001; Spector 2003). Fluorescently labeled dextran beads of 500-600 kDa are freely mobile within HeLa cell nuclei (Seksek et al. 1997; Lukacs et al. 2000), supporting the existence of an interchromosomal space through which nuclear components of a certain size can diffuse. The interchromatin compartment has been visualized indirectly by expressing green fluorescent protein (GFP)-tagged histone H2B (Kanda et al. 1998) and directly by overexpressing the intermediate filament protein vimentin, which aggregates in the interchromatin space (Bridger et al. 1998; Reichenzeller [illegible] al. 2000).

A third component of the functional nuclear architecture is the nuclear matrix, a controversial structure whose existence in intact nuclei has yet to be demonstrated (Martelli et al. 2002). The nuclear matrix is thought to support the organization of nuclear processes, such as transcription and DNA replication, and to provide a scaffold for maintaining the integrity of chromosome territories. Digestion of salt-extracted nuclei with RNase disrupts chromosome territory structure (Fey et al. 1986; Nickerson et al. 1989) and the structural integrity of chromosome territories appears to be maintained by anchoring proteins and an RNA-supported nuclear scaffold (Ma et al. 1999).

Accordingly, high salt extractions of nuclei and digestion of chromatin yield a filamentous, ribonucleoprotein-rich network that can be detected using electron microscopy (Berezney et al. 1995; Nickerson 2001). Additionally, numerous transcription factors physically interact with the nuclear matrix (see Section 1.3.2) and transcriptional activity has been shown to associate with a nuclear skeleton (Jackson et al. 1981; Jackson and Cook 1985). Nevertheless, it has proven difficult to reconcile the biochemically characterized nuclear matrix with the interchromatin compartment, which is easily observed in live and fixed cells. The Cremer group has put forward a model that satisfies proponents of both structures (Cremer et al. 2000). In it, they propose that the interchromatin compartment and the nuclear matrix are functionally and structurally equivalent, and that biochemical isolates of the nuclear matrix merely reflect the content and structure of the interchromatin space as it exists *in vivo*. In addition, the model predicts that the nuclear matrix need not be a 'fixed' scaffold; rather, the interchromosomal matrix is a dynamic entity that possesses the ability to disintegrate and reassemble in response to specific processes, such as chromosomal movement or re-ordering, or maintenance of a stable nuclear structure during terminal differentiation.

1.3.2 Subnuclear localization of transcription factors

Regulation of gene expression in eukaryotes involves a vast array of proteins and protein complexes (Lee and Young 2000; Lemon and Tjian 2000). To date, most research on eukaryotic transcription factors has used *in vitro* methods to characterize their function. These studies reveal that transcription factors are modular structures, containing separate regions for DNA binding, protein-protein interaction, and transcriptional activation or repression. It must be appreciated, however, that transcriptional regulators function in the nucleus, where they must function amid a variety of factors, including the nuclear architecture, chromatin domains, chromosome territories, and cell cycle-associated processes. Furthermore, the effects of disease mutations on the intranuclear behavior of transcription factors must be considered, as *in vitro*-based approaches do not always provide a clear understanding of the effects that mutations exert on the protein's function.

Sequence-specific transcription factors localize to various domains and compartments within the nucleus (reviewed in Corry and Underhill 2005b); several examples are discussed below. The majority of sequence-specific transcription factors form a diffuse intranuclear distribution composed of small punctate foci. Surprisingly, very few transcription factors are found in chromatin-rich areas. Instead, these proteins localize to the interchromatin compartment, possibly in association with components of the nuclear matrix. One of the best characterized families of transcription factors with regard to intranuclear compartmentalization is the nuclear steroid receptors. Early studies demonstrated that several classes of transcription factors, including nuclear receptors, were associated with the nuclear matrix (Barrack and Coffey 1980; Barrack 1983; Buttyan et al. 1983; Kaufmann et al. 1986; Alexander et al. 1987; Barrack 1987). Runt domain transcription factors (Bidwell et al. 1993; Guo et al. 1995; Merriman et al. 1995) and the pRB tumor suppressor (Mancini et al. 1994) were also demonstrated to localize to the nuclear matrix. Subsequent improvements in cell imaging techniques and the ability to create protein fusions containing fluorescent labels, such as GFP, helped provide a more definitive picture of nuclear receptor distribution. Several groups demonstrated that the glucocorticoid receptor (GR) (Htun et al. 1996; van Steensel et al. 1996), vitamin D receptor (VDR) (Barsony et al. 1997), androgen receptor (AR) (Georget et al. 1997), mineralocorticoid receptor (MR) (Fejes-Toth et al. 1998), and estrogen receptor (ER) (Htun et al. 1999) occupied numerous punctate foci evenly distributed throughout the nuclear space. Transcription factors from diverse classes display similar intranuclear distributions to these proteins, including the Wilms' Tumor transcription factor (WT1) (Englert et al. 1995; Larsson et al. 1995), the ubiquitous SP1 and SP3 proteins (He and Davie 2006), the POU domain transcription factor OCT1 (Grande et al. 1997), the *HOX* family member TLX1 (Riz et al. 2007), and the forkhead protein FOXC1 (Berry et al. 2005), among others (see also Spector 2003).

Not all transcription factors reside in the interchromatin compartment. Fluorescence imaging has revealed that the CCAAT/enhancer binding protein α (C/EBPα) localizes to centromeric heterochromatin (Tang and Lane 1999; Schaufele et al. 2001) and, upon cellular heat shock, the heat shock factor-1 (HSF1) transcription factor has been demonstrated to oligomerize in nuclear stress granules that localize to heterochromatic regions of chromosomes 9, 12, and 15 (Sarge et al. 1993; Denegri et al. 2002; Jolly et al. 2002). Interestingly, several transcription factors or portions of transcription factors localize to the nucleolus (Corsetti et al. 1995; Mancini et al. 1999; Feister et al. 2000; Rubbi and Milner 2000; Pearce et al. 2002), suggesting that it may, in addition to other functions (Olson et al. 2000), act as a "storage" compartment for certain transcription factors. Most transcription factors remain unassociated with chromosomes during cell division; however, in a process termed 'gene bookmarking,' some transcription factors, including TFIID (Christova and Oelgeschlager 2002), HSF2 (Xing et al. 2005), and RUNX2 (Young et al. 2007), bind to condensed chromosomes. This may act as an epigenetic memory mechanism that ensures efficient transcription in newly divided cells (see Sarge and Park-Sarge 2005 for review). TATA-binding protein (TBP) and its associated co-factors, as well as transcription factors associated with RNAPI have also been shown to associate with mitotic chromosomes (Chen et al. 2002; Chen et al. 2005), supporting the theory that certain gene loci may be marked for efficient re-activation in divided cells.

Another set of transcriptional regulators that interact with condensed chromatin are the so-called 'pioneering' transcription factors. These proteins bind to chromatin-embedded templates in the absence of energy-dependent chromatin remodeling and are thought to initiate early events in transcription, such as altering chromatin structure at the binding site and functioning as a platform for assembly of the transcriptional machinery. The best characterized of these pioneering factors is the FOXA1 forkhead protein, which has been shown to bind to and open compacted chromatin (Cirillo et al. 2002) and facilitate binding of ER to chromatin-embedded recognition sites (Carroll et al. 2005). The GATA4 (Cirillo et al. 2002) and PBX1 (Berkes et al. 2004) proteins act in similar manners as pioneering transcription factors. Like FOXA1 and the 'bookmarking' factors described above, PBX1 appears to mark MYOD-responsive genes by remaining stably associated with chromatin at those loci (Berkes et al. 2004). Finally, it has been shown that some transcriptional activators can overcome the silencing effects of heterochromatin by dynamically altering the local chromatin structure that contains target recognition elements (Lundgren et al. 2000; Ahmad and Henikoff 2001). Together, these observations demonstrate that compacted chromatin is not impermeable to certain DNA-binding proteins and that initial opening or remodeling of chromatin structure at promoter sequences by transcription factors is an important step in gene activation.

1.3.3 Intrinsic determinants of transcription factor subnuclear localization

Import of transcription factors into the nucleus is typically achieved through a NLS, which mediates protein entry through the nuclear pore complex (Lange et al. 2007). Once inside the nucleus, transcription factors must be targeted to their proper nuclear compartment(s). Some proteins contain signals responsible for targeting to the nucleolus (Lohrum et al. 2000) or splicing factor compartments (Eilbracht and Schmidt-Zachmann 2001); however, we have a limited understanding of analogous signals in transcription factors. While interactions with other nuclear proteins influence the intranuclear distribution of transcription factors (see Section 1.3.4), the modular nature of transcription factors has proven to be a useful tool for uncovering intrinsic determinants of subnuclear localization. As discussed below, analyses of disease-causing mutations have also provided valuable information regarding the compromised or altered compartmentalization of transcription factors in the nucleus.

Molecular dissection of the MR demonstrated that domains containing transactivation activity influence subnuclear distribution, while MR mutants compromised in DNA binding failed to form typical intranuclear clusters (Pearce et al. 2002). Together, these data suggested that MR subnuclear compartmentalization is dependent on both DNA binding and transactivation activity. Analysis of the POU-domain transcription factor PIT1 has also clarified how intrinsic determinants control intranuclear compartmentalization. In the nucleus, PIT1 displays a speckled pattern and is partitioned between a soluble fraction and a less abundant nuclear matrix-associated fraction (Mancini et al. 1999). This study also demonstrated that the POU-specific (POU_S) subdomain was capable of targeting PIT1 derivative proteins to the nuclear matrix, consistent with results showing that the DNA- (McNeil et al. 1998; Feister et al. 2000) or RNA-binding (Grondin et al. 1996) domains of other transcription factors were responsible for nuclear matrix targeting. Nuclear matrix association of PIT1 was not affected by transactivation-associated mutations in the POU_S domain or mutations in the POU-homeodomain (POU_{HD}) that abrogate DNA binding (Mancini et al. 1999), suggesting that POU_S domain-mediated gene transactivation and nuclear matrix targeting are functionally separable.

The runt domain family of transcription factors has also provided considerable information regarding subnuclear targeting. Early analyses of RUNX transcription factors revealed their association with the nuclear matrix (Bidwell et al. 1993; Merriman et al. 1995). Several years later, a 31-aa nuclear matrix-targeting sequence (NMTS) was discovered in the acute myeloid leukemia (AML)-1B, AML3, and yin-yang-1 (YY1) proteins (Zeng et al. 1997; McNeil et al. 1998; Zaidi et al. 2001). Zeng et al. (1998) also showed that co-localization of AML1B and transcriptionally active RNAPII at sites along the nuclear matrix was dependent on the NMTS, suggesting that runt domain proteins could act as a platform to which other components of the transcriptional machinery were recruited. This idea was supported by the NMTS-dependent co-localization of RUNX transcription factors with TLE family co-repressors (Javed et al. 2000). The importance of the NMTS was confirmed through mutational studies: mice carrying a deletion of the RUNX2 carboxy-terminus, including the NMTS, displayed a phenotype identical to that of null mutants (Choi et al. 2001), while NMTS point mutations disrupted interactions with SMAD partners and the yes-associated protein (YAP) (Zaidi et al. 2002; Zaidi et al. 2004). In addition, a point mutation in the AML1 NMTS was shown to disrupt granulocytic maturation of myeloid progenitor cells and caused a transformed phenotype characteristic of cells that carry an oncogenic fusion protein derived from the AML1 and ETO proteins (Vradii et al. 2005). Together, these observations provide strong evidence that interference with the subnuclear localization of runt family transcription factors contributes to their pathogenic function. Nuclear matrix targeting of the SP2 transcription factor has also been shown to be achieved by NMTS motifs in its DNA binding and transactivation domains (Moorefield et al. 2006), while the DNA-binding domains of the LHX3 and ZNF74 transcription factors (Grondin et al. 1996; Parker et al. 2000) and the combined action of the DNA-binding domain and transactivation domain of GR (Tang et al. 1998) have also been shown to possess nuclear matrix targeting capability. Together, these observations lend support to the idea that transcription factors feature distinct modules for directing their localization to nuclear compartments.

Another method of regulating transcription factor compartmentalization is alternative splicing. Larsson et al. (1995) showed that two WT1 splice variants, +KTS and –KTS (defined by the presence and absence, respectively, of a Lys-Thr-Ser tripeptide between the third and fourth zinc fingers), exhibited distinct behavior in the nucleus. The +KTS isoform co-localized with splicing factors in nuclear speckles or clusters, while the –KTS isoform displayed a predominantly diffuse pattern. Likewise, splice forms of the ER and the zinc finger protein NP/NMP4 were shown to localize to distinct subnuclear compartments (Feister et al. 2000; Price et al. 2001). Finally, another study of WT1 showed that its subnuclear localization is independent of DNA binding activity and that truncations of the protein caused localization to nuclear speckles and conferred an ability to recruit wild type –KTS WT1 to speckle domains (Englert et al. 1995).

1.3.4 Regulation of transcription factor subnuclear localization through co-factor interactions

Interactions between transcription factors and other nuclear proteins or components presents an attractive method of regulating gene expression by spatially and temporally sequestering binding partners to distinct nuclear compartments until conditions are appropriate for transcription. A similar system of regulation exists for transcription factors that exhibit signal-induced nucleocytoplasmic shuttling, for example, the steroid receptors (DeFranco 1999; Kaffman and O'Shea 1999). Intranuclear sequestering, however, represents a more complex scenario, since the proteins are already present in the nucleus and therefore must be kept separate without the aid of a physical barrier (i.e., the nuclear envelope). Several examples, discussed below, describe the intranuclear distribution of transcription factors and their interacting co-regulators, and demonstrate that sequestration of one or both proteins plays a role in transcriptional control.

Expression of nuclear receptor target genes is regulated through interactions with co-factors (McKenna et al. 1999). AR interacts with members of the SRC/p160 family of co-activators, including the steroid receptor coactivator 1 (SRC1) and GR-interacting protein 1 (GRIP1) (Black and Paschal 2004). SRC1 and GRIP1 occupy distinct subnuclear foci (Nazareth et al. 1999; Baumann et al. 2001) and, in the presence of agonist-bound AR, are recruited to AR domains (Nazareth et al. 1999; Black et al. 2004). A similar scenario has been described for GR and the receptor-interacting protein 140 (RIP140) co-repressor (Tazawa et al. 2003). Anti-androgen treatment disrupts the interaction between AR and GRIP1 (Karvonen et al. 2002), while mutations in its DNA- or ligand-binding domains cause AR to accumulate in SRC1 or GRIP1 domains (Nazareth et al. 1999; Black et al. 2004). Furthermore, Black et al. (2004) demonstrated that the CREB-binding protein (CBP), which is normally recruited by p160 family proteins to AR-responsive genes, mislocalized to GRIP1 foci in the presence of AR that possessed mutations associated with androgen insensitivity syndrome.

Studies of PIT1 have also provided information on the intranuclear relationship between transcription factors and co-regulatory proteins. The PIT1-interacting transcription factor C/EBPα localizes to centromeric heterochromatin (see Section 1.3.2), but when co-expressed with PIT1, relocates from these domains to regions occupied by PIT1 (Enwright et al. 2003). Mutations associated with DNA binding and transactivation defects have been shown to cause co-localization of PIT1 and C/EBPα in heterochromatic regions (Day et al. 2003; Enwright et al. 2003). In addition, a C/EBPα protein lacking DNA binding ability failed to localize to the pericentromeric domains and did not associate with nuclear PIT1 (Day et al. 2003). Recently, PIT1 was shown to redistribute the nuclear receptor co-repressor (N-CoR) and silencing mediator of retinoid and thyroid hormone receptors (SMRT) proteins from nuclear foci to PIT1-occupied domains (Voss et al. 2005). However, a disease-causing mutation in the PIT1 POU$_{HD}$ abrogated co-repressor redistribution and triggered PIT1 redistribution to SMRT foci.

Overall, the intranuclear interactions between wild type or mutant transcription factors and their co-regulatory partners demonstrate that intranuclear compartmentalization can be altered by mutations in DNA-binding domains (Day et al. 2003; O'Toole et al. 2003; Black et al. 2004), transactivation/repression domains (Choi et al. 2001; Francastel et al. 2001; Yin et al. 2001; Rajaram and Kerppola 2004), or both (Mancini et al. 1999; Pearce et al. 2002). These observations imply that determinants of subnuclear localization vary among different transcription factor classes and also among proteins within these classes.

1.3.5 Intranuclear dynamics

The ability to fuse fluorescent moieties onto nuclear proteins and nucleic acid molecules affords researchers an opportunity to examine the dynamics of nuclear components in live cells. Early studies using photobleaching demonstrated that fluorescently labeled nuclear proteins exhibited high rates of mobility (see Dundr and Misteli 2001 for a review). It was also found that, in most cases, the movement of nuclear constituents occurs through passive diffusion, rather than an energy-dependent mechanism (Misteli 2001), suggesting that components of the nucleus interact through stochastic, transient associations. While seemingly inefficient, this behavior is economical from the nucleus' perspective, since passive diffusion requires no energy source and ensures nuclear components are constantly in flux, ready to respond to external stimuli. In addition, stochastic interactions are consistent with the relatively slow and inefficient assembly of transcription complexes *in vitro* (Kadonaga 1990; Bral et al. 1998). A model predicting transient interactions between freely diffusing proteins and relatively immobile templates, such as chromatin, is also consistent with observations of protein mobility in live cells which, based on molecular weight, tends to be slower than anticipated (Misteli 2001).

Results from live cell imaging studies suggest that chromosome territories are generally immobile in the nucleus (reviewed in (Dundr and Misteli 2001; Lanctot et al. 2007). Nevertheless, several studies demonstrate that chromosomes in metazoan cells undergo constrained diffusion within a given territory (Marshall et al. 1997; Vazquez et al. 2001), while long-range, actin-dependent chromosomal movement has also been proposed (Chuang et al. 2006; Dundr et al. 2007). Several groups have employed artificial tandem arrays of regulatory binding elements to address the intranuclear mobility of chromatin-embedded loci (Robinett et al. 1996; Muller et al. 2001). Studies using these arrays have demonstrated chromatin decondensation in the presence and absence of transcription (Tumbar et al. 1999; Muller et al. 2001), as well as mobilization of the locus from the nuclear periphery in response to DNA replication and gene activation (Li et al. 1998; Tumbar and Belmont 2001). In contrast to the largely constrained motion of chromosome domains, individual chromatin loci have been shown to exhibit fast, short-range motion (Chubb et al. 2002; Levi et al. 2005) and these 'jumps' appear to be dependent on an energy source (Levi et al. 2005).

Several studies have shown that cell type and differentiation status influence chromatin dynamics. In particular, several loci involved in the immune response are dynamically repositioned in nucleus, depending on the stage of cellular development or differentiation (Brown et al. 2001; Kosak et al. 2002; Hewitt et al. 2004; Ragoczy et al. 2006). In addition, the *Mash1* locus, which is involved in cellular commitment to a neural fate, undergoes chromatin reorganization and intranuclear repositioning upon neural induction (Williams et al. 2006), while genes involved in muscle differentiation were shown to be localized proximal to nuclear RNA-splicing domains in post-mitotic muscle cells (Moen et al. 2004). Upon commitment to a specific differentiation pathway, pluripotent embryonic stem cells undergo global nuclear reorganization, including formation of defined chromatin territories (reviewed in Meshorer and Misteli 2006) and repositioning of genes responsible for maintaining a pluripotent state (Wiblin et al. 2005). Together, these observations indicate that chromatin-embedded loci are highly dynamic and that reorganization of chromatin domains is a central event in cellular differentiation. Nonetheless, observations of endogenous loci demonstrate that there is no straightforward correlation between the subnuclear positioning of a locus and its transcriptional status. For example, the *CFTR* (Zink et al. 2004), *β-globin* (Ragoczy et al. 2006), *IgH* (Kosak et al. 2002), and *Mash1* (Williams et al. 2006) loci all require localization away from the nuclear periphery for optimum expression, while some genes in both yeast (Brickner and Walter 2004; Casolari et al. 2004; Casolari et al. 2005) and mammalian (Hewitt et al. 2004) cells remain transcriptionally active in the nuclear periphery. In the latter cases, it is hypothesized that peripheral localization of active gene loci brings them in the vicinity of nuclear pores, facilitating RNA export to the cytoplasm (see Akhtar and Gasser 2007 for a recent review).

With the exception of the core histones, chromatin-associated proteins, including transcription factors and architectural proteins, exhibit high rates of mobility within the nucleus (Misteli 2001). In general, these proteins, exemplified by histone H1 (Lever et al. 2000; Misteli et al. 2000), HMG17 (Phair and Misteli 2000), and HP1 (Cheutin et al.

2003; Festenstein et al. 2003), dynamically interact with chromatin loci in a transient, "hit-and-run" manner. Sequence-specific transcription factors also show high rates of intranuclear mobility. Live cell imaging of the GR demonstrated rapid exchange on and off a chromatinized mouse mammary tumor virus (MMTV) promoter array (McNally et al. 2000), while photobleaching experiments showed variable mobility of ER, depending on the presence or absence of agonists and antagonists (Stenoien et al. 2001). The same researchers demonstrated that co-activators for GR (Becker et al. 2002) and ER (Stenoien et al. 2001) exhibited similar rates of mobility similar to their steroid receptor partners. The Hager group later investigated real-time GR interaction with the MMTV array using an efficient UV-crosslinking procedure and demonstrated rapid, transient binding of GR to the template followed by targeted recruitment of the ATP-dependent SWI/SNF chromatin remodeling complex (Nagaich et al. 2004). These observations confirm that transcription factor interaction with chromatin-embedded templates is highly dynamic. Other transcription factors show high turnover rates and interact transiently with chromatin templates (Phair et al. 2004; Bosisio et al. 2006), similar to steroid receptors. Additionally, several studies have indicated that molecular chaperones play a role in the intranuclear mobility of steroid receptors in an energy-dependent manner (Elbi et al. 2004; Agresti et al. 2005). These results coincide with data showing structural alterations in nuclear receptors upon co-factor interaction or DNA binding (Kumar and Thompson 2003; Kumar et al. 2004), suggesting that some transcription factors require chaperone-mediated refolding in order to move from target to target in the nucleus (reviewed in Hager et al. 2006). Finally, studies of mutant transcription factors suggest that DNA binding activity is an important determinant of intranuclear mobility (Schaaf and

Cidlowski 2003; Farla et al. 2004; Karpova et al. 2004; Schaaf et al. 2006). For example, analyses of wild type and mutant forms of GR and the NF-κB subunit p65 found that high mobility correlated with low DNA binding affinity and a more random distribution in the nucleus (Schaaf et al. 2005; Schaaf et al. 2006). Together, these results demonstrate that the dynamic nature of transcriptional regulatory proteins is an important determinant of their function and that the coordinated and dynamic interaction between the transcriptional machinery and target chromosomal loci is critical for proper regulation of eukaryotic gene expression (reviewed in Schneider and Grosschedl 2007).

1.3.6 Functional organization of eukaryotic transcription

Collectively, observations of gene expression in the nucleus form a picture of a highly dynamic process that is organized at multiple levels (van Driel et al. 2003). Initial studies of the intranuclear localization of transcription demonstrated that active RNAPII coalesces in 1000–2000 punctate foci within the nucleus (Jackson et al. 1993; Wansink et al. 1993). Concurrently, the van Driel group showed that similar foci, both in number and appearance, were formed by the GR (van Steensel et al. 1995) and MR (van Steensel et al. 1996). Surprisingly, the majority of GR or MR foci did not co-localize with RNAPII or splicing factors, although GR and MR co-localized with each other to a large extent. These findings suggested that a large proportion of a given transcription factor might reside in "storage" domains that are not associated with transcriptionally active genes. Other sequence-specific transcription factors, including OCT1, E2F, and the GATA1, -2, and -3 proteins, were also shown to form punctate nuclear foci that displayed limited co-localization with RNAPII/transcription domains (Elefanty et al. 1996; Grande et al. 1997). OCT1 was later shown to co-localize with the PSE-binding transcription factor (PTF), RNAPII, TBP, and SP1 proteins in transcriptionally active foci termed OPT domains (Pombo et al. 1998). Together, these observations lend support to the "transcription factory" model of gene expression (Pombo et al. 2000; Martin and Pombo 2003), which proposes that RNA synthesis and processing occur in discrete subnuclear foci that contain factors required for transcription of a particular gene or set of genes. Subsequent studies confirmed that transcription occurs in distinct foci where active genes from non-contiguous regions of a chromosome (and other chromosomes) localize to preassembled 'factories' containing transcription and splicing machinery (Osborne et al. 2004).

The functional organization of transcription factories results from the inherently dynamic nature of their components. Individual gene loci on a chromatin fiber are highly mobile (see Section 1.3.5), as is the majority of RNAPII, which possesses a transiently immobile fraction that represents actively transcribing polymerase (Kimura et al. 2002). Splicing factors are dispersed from large interchromatin granules to transcription factories upon transcriptional activation (Spector 1996), implying that transcription and mRNA processing are spatiotemporally linked *in vivo* (reviewed in Bentley 2005). In the current model of nuclear transcriptional organization, which is supported by an extensive body of data (see Cremer et al. 2004 and references therein), transcription and splicing occur at perichromatin fibrils. These structures correspond to nascent mRNA transcripts that are associated with transcription and splicing machinery (Fakan 1994; Cmarko et al. 1999). Perichromatin fibrils are found in the perichromatin region, a functional domain formed at the interface between chromosome territories and the interchromatin

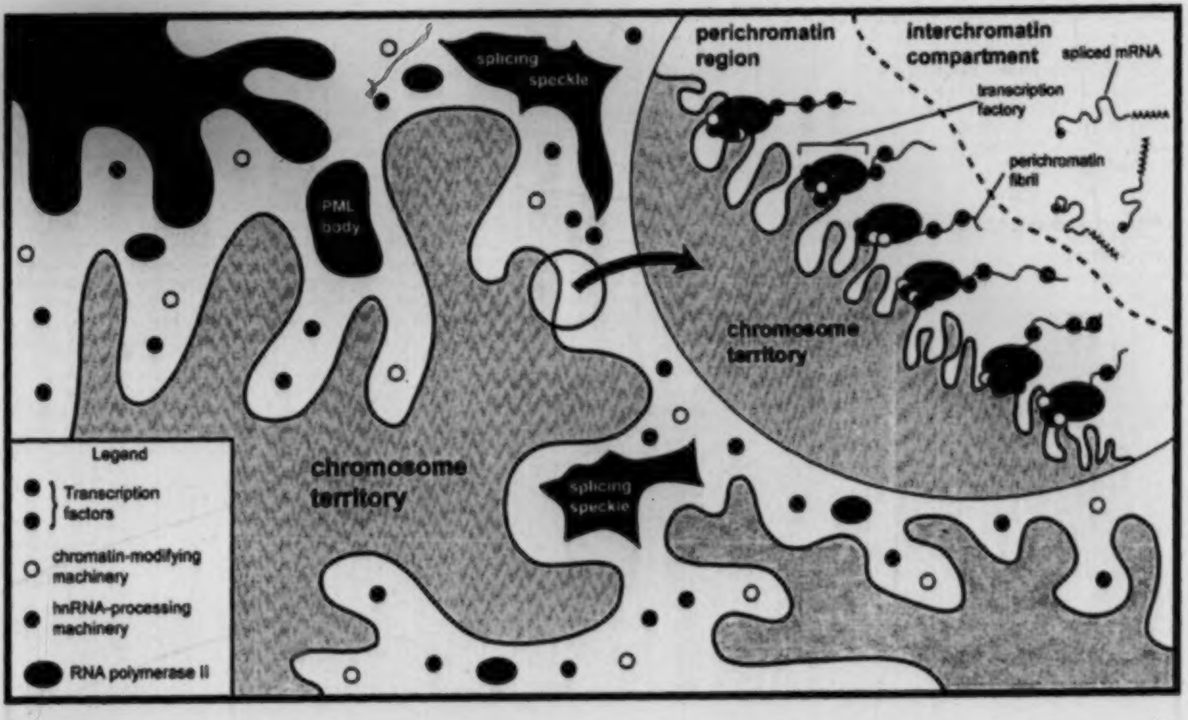

perichromatin region
interchromatin compartment
spliced mRNA
transcription factory
perichromatin fibril
chromosome territory
splicing speckle
PML body
splicing speckle
chromosome territory
Legend
Transcription factors
chromatin-modifying machinery
hnRNA-processing machinery
RNA polymerase II

Figure 1-5. Organization of eukaryotic gene expression. The chromosome territory-interchromatin compartment model of nuclear organization (see Cremer and Cremer 2001; Cremer et al. 2006) postulates that chromosomes occupy distinct territories in the nucleus (indicated by green, blue, and pink domains), while the interchromatin compartment (grey) accommodates components of transcription and DNA replication and repair. Transcription occurs at discrete 'factories' in the perichromatin region (inset). Nascent mRNA, in the form of perichromatin fibrils, is co-transcriptionally spliced by factors that mobilize from splicing speckles (red). Polyadenylated, 5'-capped mRNA is released into the interchromatin compartment and is transported to the cytoplasm for translation. Sequence-specific transcription factors, chromatin- and hnRNA-modifying machinery, and RNA polymerase II (see legend) are freely mobile in the interchromatin compartment; other nuclear domains, including PML bodies (purple) are located in the interchromatin compartment. See Section 1.3.6 for details.

compartment (Fakan 2004). The model predicts that chromatin fibers containing active or potentially active genes are dynamically looped out of chromosome territories, combining with components of the transcriptional machinery that have been mobilized from domains in the interchromatin compartment to the perichromatin region (see Fig. 1-5). Thus, eukaryotic gene expression is a paradoxical process, in which the steady-state organization of transcription is achieved through the continuous flux and mobility of the transcriptional machinery and its chromatin substrates.

1.4 Objective

PAX3 is a well-characterized transcription factor that initiates key developmental pathways, including myogenesis, melanogenesis, and neurogenesis. Mutations and deletions of *Pax3* and *PAX3* lead to the *Sp* phenotype in mice and WS in humans, respectively. These disorders are characterized by the failure of PAX3-specified cell populations to proliferate and differentiate, leading to specific defects in pigmentation, limb musculature, and craniofacial development. *In vitro* studies have provided much of our knowledge regarding the functional behavior of PAX3, including DNA binding and target gene regulation, and effects of disease-causing mutations on these properties. However, there remains a gap in knowledge between our understanding of PAX3 function at a molecular level and how improper PAX3 behavior leads to a disease phenotype at the tissue/organ level. Furthermore, correlating *PAX3* mutations with their associated phenotypes has proved an elusive task, suggesting that more information is required to determine how each mutation contributes to a PAX3-related disorder.

Finally, the nature of the determinants that underlie PAX3 function and behavior, particularly at the *in vivo* level, is poorly understood.

The aim of this study is to elucidate, through molecular and cellular analyses, the determinants responsible for controlling PAX3 behavior, with a major focus on how disease mutations affect PAX3 behavior. Here, we establish how PAX3 interacts with various target sites, including a subset of putative target promoter elements, and characterize how functional cooperativity and the differential use of the two PAX3 DNA-binding domains permits recognition of disparate target sequences. We also demonstrate that disease-causing mutations in the PD and HD exert variable effects on the DNA binding and transactivation behavior of PAX3. Following this, we examine the behavior of PAX3 from an *in vivo* perspective, analyzing its subnuclear localization and intranuclear mobility, as well as intrinsic determinants that influence these traits. Also investigated are the consequences of disease mutations on PAX3 behavior in the nucleus and their effects on compartmentalization and mobility. Collectively, these studies provide an entry point to understanding the various intrinsic factors that control PAX3 behavior, and shed light on how disease mutations interfere with this mechanism of regulation.

CHAPTER 2. METHODS AND MATERIALS

2.1 Plasmid construction

2.1.1 *Bacterial expression constructs*

Hexahistidine-tagged proteins were made by inserting PCR-amplified *Pax3* cDNA fragments into the pET21a expression plasmid (Novagen). A *Pax3c* Q- template (see Section 1.1.2) was used for all amplifications. PCR reactions were performed in a 20 μL volume containing 10X *Pfu* buffer (10 mM KCl, 20 mM Tris-HCl (pH 8.8), 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100), 1 mM dNTPs, 1 unit *Pfu* Turbo polymerase (Stratagene), and 2 μM each primer (primer sequences are listed in Table 1) for 35 cycles of amplification under the following conditions: denaturation at 94°C for 1 min, 30 s primer annealing at the appropriate temperature, extension at 72°C, followed by an additional extension period of 7 min at 72°C. All PCR fragments were created to contain restriction sites on the ends that would allow for insertion in-frame with the carboxy-terminal His-Tag encoded by pET21a. The vector also encodes a T7 epitope tag that is fused to the amino-terminus of the protein. After restriction digestion, the fragments were ligated into pET21a using T4 DNA ligase (Invitrogen). PAX3-derived constructs in pET21a are listed in Table 2.

Recombinant PDs from PAX1, PAX2, PAX6, and PAX7, and HDs from PAX6, PAX7, PRRX1, PITX2, MSX1, and the rudimentary HD from PAX2 were constructed as described above for PAX3-derived constructs and expressed from pET21a and are detailed in Table 3. Chimeric HDs containing the first two α-helices of the PRRX1 or PITX2 HDs and the recognition helix of the PAX3 HD were created by PCR amplifying the sequence containing the amino-terminal extension and helices 1 and 2 from the pET21a-PRRX1 HD (pET.PRRX1HD.F and pET.PRRX1HD.chimera.R) or pET21a-

Table 1. Primers used for PCR amplification of inserts for expression constructs.

Name	Sequence (5'-3')	Restriction site
pET.PAX3PD.F	CAGGATCCGGCCAGGGCCGAGTCAA	*Bam*HI
pET.PAX3PD.R	CCAGAAGCTTTTTTCCAAATTTACTCC	*Hin*dIII
pET.PAX3HDΔQRR.F	CAGGATCCAGCAGAACCACCTTCAC	*Bam*HI
pET.PAX3HD+NLS.F	CAGGATCCAGCTCTGAACCTGATTTAC	*Bam*HI
pET.PAX3HD.R	CCTGCGGCCGCAAGCTTTCCAGCTTGTTTCCTCC	*Hin*dIII/*Eag*I
pET.PAX3HD.F	ACGGATCCCAGCGCAGGAGCAGAACC	*Bam*HI
pET.PAX3linker.R	CCTGCGGCCGCAAGCTTCCTCTTCAGCGGTAA	*Eag*I/*Hin*dIII
pET.PAX3linker.F	CAGGATCCAAAGGAGAAGAGGAGGAGGCG	*Bam*HI
pET.PAX3PD-Oct.R	TACAAGCTTTTTAGCCTTTTTCTCGCT	*Hin*dIII
pET.HD.Ntermdel.F	CAGGATCCTTCACGGCAGAGCAGCTG	*Bam*HI
pET.PAX3HD.helix1-2.R	ACAAAGCTTCGCCCTCTGGGCCA	*Hin*dIII
pET.PAX3HD+Oct.F	ACGGATCCGCTAAACACAGCATCGAT	*Bam*HI
pET.PAX3HD-Oct.F	ACGGATCCGCCTCTGCACCTCAGTCA	*Bam*HI

pET.PAX1PD.F	CCTAGATCTACGTACGGCGAAGTGAAC	Bg*l*II
pET.PAX1PD.R	CCTAAGCTTAATCTTATTTCGCAGGATG	*Hind*III
pET.PAX2PD.F	CCTGGATCCGGGCACGGGGGTGTGAAC	*Bam*HI
pET.PAX2PD.R	CCTAAGCTTAACTTTGGTCCGGATGATTC	*Hind*III
pET.PAX6PD.F	CCTGGATCCAGTCACAGCGGAGTGAATC	*Bam*HI
pET.PAX6PD.R	CCTAAGCTTAGCCAGGTTGCGAAGAAC	*Hind*III
pET.PAX7PD.F	CAGGATCCGGCCAAGGCCGGGTCAAT	*Bam*HI
pET.PAX7PD.R	GCAAAGCTTCTTCCCGAACTTGATTCT	*Hind*III
pET.PRRX1HD.F	CAGGATCCCAGCGAAGGAATAGGA	*Bam*HI
pET.PRRX1HD.R	CCAGAAGCTTTCTCTCATTCCTGCGGA	*Hind*III
pET.PRRX1HD.chimera.R	ACAAGCTTCACCCGGCGGGCAAGGTCTTCT	*Hind*III
pET.PITX2HD.F	CAGGATCCCAAAGGCGGCAGCGGACT	*Bam*HI
pET.PITX2HD.R	CCAGAAGCTTGCGCTCCCTCTTTCTCC	*Hind*III
pET.PITX2HD.chimera.R	ACAAGCTTGGTCCACACAGCGATTTCTTC	*Hind*III
pET.PAX2HD.F	CAGGATCCCACTTGCGAGCTGACACCTTCAC	*Bam*HI

pET.PAX2HD.R	CCTGCGGCCGCAAGCTTTGCAGATAGACTCGACTTG	*Hin*dIII
pET.PAX6HD.F	CAGGATCCCTGCAAAGAAATAGAACATCCTTTACC	*Bam*HI
pET.PAX6HD.R	CCGCAAGCTTTTTTTCTTCTCTTCTCCATTTGG	*Hin*dIII
pET.PAX7HD.F	CAGGATCCCAGCGACGCAGTCGGACC	*Bam*HI
pET.PAX7HD.R	GCAAAGCTTTCCTGCCTGCTTACGCCA	*Hin*dIII
pET.MSX1HD.F	CAGGATCCCCAGCCTGCACCCTCCGC	*Bam*HI
pET.MSX1HD.R	ACTAAGCTTCAGCTCTGCCTCTTGTAGTC	*Hin*dIII
GFP.PAX3PD.F	ACGAATTCACGGCCAGGGCCGAGTCAAC	*Eco*RI
GFP.PAX3PD.R	ACGGATCCACTTTCCAAATTTACTCCT	*Bam*HI
GFP.PAX3.F	CAGAATTCTGATGACCACGCTGGCC	*Eco*RI
GFP.PAX3.R	CAGGATCCTGGAACGTCCAAGGCTT	*Bam*HI
GFP.PAX3linker.F	CAGAATTCTAGCCGCCACCATGAAAGGAGAAGAGGAGGA	*Eco*RI
GFP.PAX3HD.F	CCGAATTCATGACCTCTGAACCTGATTTACC	*Eco*RI
GFP.PAX3HD.R	ACCGGATCCTGATTGGCTCCAGCTTG	*Bam*HI
GFP.PAX3HD+Oct.F	ACGAATTCACATGGCTAAACACAGCATCGAT	*Eco*RI

GFP.PAX3HD-Oct.F	ACGAATTCACATGGCCTCTGCACCTCAGTCA	*Eco*RI
PRRX1.FL.top	ATGACCTCCAGCTACGGGCACGTTCTGGAGCGGCAACCGGCGCTGGGCGGCCGCTTGGACAGC	
PRRX1.FL.bottom	CCGGGCTGTCCAAGCGGCCGCCCAGCGCCGGTTGCCGCTCCAGAACGTGCCCGTAGCTGGAGGTCAT	
GFP.PRRX1.F	CAGAATTCACATGACCTCCAGCTAC	*Eco*RI
GFP.PRRX1.R	CAGGATCCACGAATCCGTTATGAAG	*Bam*HI
GFP.PRRX1HD.F	CCGAATTCATGACCCTGAACTCAGAAGAAAAAA	*Eco*RI
GFP.PRRX1HD.R	ACCGGATCCAGCATGGCTCTCTCATTC	*Bam*HI
GFP.PITX2HD.F	CCGAATTCATGACCGGCGCCGAGGACCCGTCT	*Eco*RI
GFP.PITX2HD.F	ACCGGATCCTGCTGGTTGCGCTCCC	*Bam*HI
pCI.TLE4.F	CTGAATTCATGATTCGCGACCTGAGC	*Eco*RI
pCI.TLE4.R	CTGGTACCTTAATAAATAACTTCATAAACTGTGG	*Kpn*I

Primers are shown in 5'-3' direction; restriction sites used to ligate PCR-amplified inserts into recipient vector are also given.

Table 2. PAX3 expression constructs in pET21a.

Construct	Primer set	Amino acids
pET21a-PD	pET.PAX3PD.F and pET.PAX3PD.R	34-164
pET21a-PD-Oct	pET.PAX3PD.F and pET.PAX3PD-Oct.R	34-185
pET21a-PD+linker	pET.PAX3PD.F and pET.PAX3linker.R	34-218
pET21a-PDHD	pETPAX3PD.F and pET.PAX3HD.R	34-279
pET21a-HD+linker	pET.PAX3linker.F and pET.PAX3HD.R	164-279
pET21a-HD+Oct	pET.PAX3HD+Oct.F and pET.PAX3HD.R	184-279
pET21a-HD-Oct	pET.PAX3HD-Oct.F and pET.PAX3HD.R	196-279
pET21a-HD+NLS	pET.PAX3HD+NLS.F and pET.PAX3HD.R	209-279
pET21a-HD	pET.PAX3HD.F and pET.PAX3HD.R	219-279
pET21a-HDΔh3	pET.PAX3HD.F and pET.PAX3HD.helix1-2.R	219-256
pET21a-HD HDΔQRR	pET.PAX3HDΔQRR.F and pET.PAX3HD.R	222-279
pET21a-HDΔN	pET.HD.Ntermdel.F and pET.PAX3HD.R	226-279

See Table 1 for primer sequences; amino acid positions are relative to the full length PAX3 protein.

Table 3. Recombinant paired domain and homeodomain expression constructs in pET21a.

Construct	Primer set	Amino acids
pET21a-PAX1 PD	pET.PAX1PD.F and pET.PAX1PD.R	4-131
pET21a-PAX2 PD	pET.PAX2PD.F and pET.PAX2PD.R	16-143
pET21a-PAX6 PD	pET.PAX6PD.F and pET.PAX6PD.R	4-145
pET21a-PAX7 PD	pET.PAX7PD.F and pET.PAX7PD.R	34-165
pET21a PAX2 'HD'	pET.PAX2HD.F and pET.PAX2HD.R	227-286
pET21a-PAX6 HD	pET.PAX6HD.F and pET.PAX6HD.R	210-270
pET21a-PAX7 HD	pET.PAX7HD.F and pET.PAX7HD.R	215-275
pET21a-PRRX1 HD	pET.PRRX1HD.F and pET.PRRX1HD.R	94-154
pET21a-PITX2 HD	pET.PITX2HD.F and pET.PITX2HD.R	92-152
pET21a-MSX1 HD	pET.MSX1HD.F and pET.MSX1HD.R	162-255

See Table 1 for primer sequences; amino acid positions are relative to the full length murine (PAX1 PD, PAX2 PD, PAX6 PD) and human (PAX7 PD, PAX2 'HD', PAX6 HD, PAX7 HD, PRRX1 HD, PITX2 HD, MSX1 HD) proteins.

PITX2 HD (pET.PITX2HD.F and pET.PITX2HD.chimera.R) constructs and ligating into wild type or residue 50-mutant (see below) pET21a-PAX3 HD cut with *Bam*HI and *Hin*dIII. The resulting constructs are pET21a-PRRX1-PAX3 HD and pET21a-PITX2-PAX3 HD. All constructs were verified by sequencing.

2.1.2 *Mammalian expression constructs*

Two constructs were used to express the untagged full length PAX3c isoform: pMT2-PAX3 (described in Underhill et al. 1995) and pcDNA3.1-PAX3 (described in Corry and Underhill 2005a). An *Eco*RI-*Bam*HI PCR fragment containing the entire PAX3 coding sequence was inserted into pEGFP-N1 (Clontech) to allow expression of PAX3 with a carboxy-terminal GFP tag; primers used were: GFP.PAX3.F and GFP.PAX3.R. PAX3 was also expressed from pEGFP-C1 (Clontech) (GFPC1.PAX3.F and GFPC1.PAX3.R), which fuses GFP to the amino-terminus. The PRRX1-GFP construct was created by first ligating oligonucleotides encoding the amino-terminal portion of PRRX1 (PRRX1.FL.top and PRRX1.FL.bottom) into the pCGNphox construct (previously described in Grueneberg et al. 1995) to create the full length PRRX1 sequence. Then, the full PRRX1-coding sequence was amplified using the GFP.PRRX1.F and GFP.PRRX1.R primers and ligated into pEGFP-N1. The pCI-HA-PITX2 construct has been described previously (Berry et al. 2006). Hemagglutinin (HA)-tagged TLE4 was obtained by PCR amplification of cDNA encoding the TLE4a (Invitrogen) and TLE4b and TLE4c (Open Biosystems) isoforms with the pCI.TLE4.F and pCI.TLE4.R primers and ligating into the pCI vector (Promega) carrying an amino-terminal HA-encoding sequence.

Regions of PAX3 were PCR amplified and inserted into pEGFP-N1 to analyze determinants of PAX3 subnuclear localization and mobility. These constructs are listed in Table 4. Fluorescently tagged versions of the PRRX1 and PITX2 HDs were created by amplifying the HD-encoding sequence and adjacent NLS from the respective full length constructs and ligating the fragment into pEGFP-N1. Primers used were GFP.PRRX1HD.F and GFP.PRRX1HD.R for pEGFP-N1-PRRX1 HD, and GFP.PITX2HD.F and GFP.PITX2HD.R for pEGFP-N1-PITX2 HD. Fluorescently labeled chimeric HDs, containing the first two α-helices of the PRRX1 or PITX2 HDs and the recognition helix of the PAX3 HD, were also constructed: pEGFP-N1-PRRX1-PAX3 HD was created by amplifying the coding sequence from pET21a-PRRX1-PAX3 HD using the GFP.PRRX1HD.F and GFP.PAX3HD.R primers, and pEGFP-N1-PITX2-PAX3 HD was constructed by amplifying the coding sequence from pET21a-PITX2-PAX3 HD using the GFP.PITX2HD.F and GFP.PAX3HD.R primers.

2.1.3 *Luciferase reporter constructs*

A 477 bp portion of the *MITF* promoter that spans the region from –382 to +95 relative to the transcription start site was PCR amplified from human genomic DNA using primers MITF.F.Kpn (5'-CCGGTACCGTCGGAAGTGGCAGTTA-3') and MITF.R.Bgl (5'-CCAGATCTCTTATCCCTCCCTCTAC-3'). The purified fragment was restriction digested with *Kpn*I and *Bgl*II, and ligated into *Kpn*I and *Bam*HI-digested pGL3-basic (Promega) to generate *MITF-Luc*. The *Trp1-Luc* construct has been described previously (Galibert et al. 1999).

Table 4. PAX3 expression constructs in pEGFP-N1.

Construct	Primer set	Amino acids
pEGFP-N1-N-PDHD	GFP.PAX3.F and GFP.PAX3HD.R	1-279
pEGFP-N1-PD	GFP.PAX3PD.F and GFP.PAX3PD.R	34-164
pEGFP-N1-PDHD	GFP.PAX3PD.F and GFP.PAX3HD.R	34-279
pEGFP-N1-linker-HD	GFP.PAX3linker.F and GFP.PAX3HD.R	164-279
pEGFP-N1-linker-C	GFP.PAX3linker.F and GFP.PAX3.R	164-479
pEGFP-N1-HD+Oct	GFP.PAX3HD+Oct.F and GFP.PAX3HD.R	184-279
pEGFP-N1-HD-Oct	GFP.PAX3HD-Oct.F and GFP.PAX3HD.R	196-279
pEGFP-N1-HD	GFP.PAX3HD.F and GFP.PAX3HD.R	209-279
pEGFP-N1-HD-C	GFP.PAX3HD.F and GFP.PAX3.R	209-479

See Table 1 for primer sequences; amino acid positions are relative to the full length PAX3 protein.

2.2 Mutant expression constructs

The pcDNA3.1-PAX3 *Sp*^{d} construct has been described previously (Corry and Underhill 2005a). Other mutant PAX3 expression plasmids were created using codon-specific primers (see Table 5) and the QuikChange Site-Directed Mutagenesis Kit (Stratagene) and the pcDNA3.1-PAX3 template, except for the Y90H, A196T, and Q391H mutants, which used pEGFP-N1-PAX3 as the template. To create recombinant mutant PD-, HD-, or PDHD-expressing constructs, we performed PCR amplification of the coding region using the full length mutant template and inserted the fragment into pET21a. pEGFP-N1-PAX3 constructs carrying PD mutations were created by excising an *Xma*I fragment from wild type pEGFP-N1-PAX3 and exchanging it with the corresponding fragment from the mutant template; mutant HD constructs in pEGFP-N1 were created by exchanging a *Kpn*I fragment from the wild type sequence for one containing the V265F or R271G mutation.

Constructs to express PAX3, PRRX1, and PITX2 HD proteins containing residue 50 mutations were created using mutagenic primers and the QuikChange Kit (see above),. In each case, the wild type pET21a construct was used as a template. GFP-tagged versions were created by PCR amplification of the coding sequence and inserting into pEGFP-N1. The presence of the correct mutation was verified by sequencing.

2.3 Hexahistidine-tagged protein expression and purification

pET21a plasmids were transformed into BL21DE3 *E. coli* and used for protein expression. Briefly, 1250 μL of overnight culture was used to inoculate 50 mL LB

Table 5. Primers for site-directed mutagenesis.

Name	Sequence (5'-3')
PAX3.F45L.top	CCAGCTCGGAGGAGTACTTATCAAAGGCAGG
PAX3.F45L.bottom	CCTGCCGTTGATAAGTACTCCTCCGAGCTGG
PAX3.N47H.top	GAGGAGTATTTATCCACGGCAGGCCTCTGCC
PAX3.N47H.bottom	GGCAGAGGCCTGCCGTGGATAAATACTCCTC
PAX3.G81A.top	CGCGTGTCCCATGCTTGCGTCTCTAAGATCC
PAX3.G81A.bottom	GGATCTTAGAGACGCAAGCATGGGACACGCG
PAX3.S84F.top	CCCATGGTTGCGTCTTTAAGATCCTGTGCAGG
PAX3.S84F.bottom	CCTGCACAGGATCTTAAAGACGCAACCATGGG
PAX3.Y90H.top	AGGATCCTGTGCAGGCACCAGGAGACAGGCTCC
PAX3.Y90H.bottom	GGAGCCTGTCTCCTGGTGCCTGCACAGGATCCT
PAX3.A196T.top	ATCCTGAGTGAGCGAACCTCTGCACCTCAGTCA
PAX3.A196T.bottom	TGACTGAGGTGCAGAGGTTCGCTCACTCAGGAT
PAX3.V265F.top	AGGCCCGAGTGCAGTTCTGGTTTAGCAACCG
PAX3.V265F.bottom	CGGTTGCTAAACCAGAACTGCACTCGGGCCT
PAX3.R271G.top	GGTTTAGCAACCGCGGTGCAAGATGGAGGAA
PAX3.R271G.bottom	TTCCTCCATCTTGCACCGCGGTTGCTAAACC
PAX3.Q391H.top	GGCAATGGCCTTTCACCTCACGTAATGGGACT
PAX3.Q391H.bottom	AAGTCCCATTACGTGAGGTGAAAGGCCATTGCC
PAX3HD.S50Q.top	CAGGTCTGGTTTCAGAACCGCCGTGC
PAX3HD.S50Q.bottom	GCACGGCGGTTCTGAAACCAGACCTG
PAX3HD.S50K.top	CAGGTCTGGTTTAAGAACCGCCGTGC

PAX3HD.S50K.bottom	GCACGGCGGTTCTTAAACCAGACCTG
PRRX1HD.Q50S.top	GCGAGAGTGCAGGTGTGGTTTAGCAACCGAAGAGCC
PRRX1HD.Q50S.bottom	GGCTCTTCGGTTGCTAAACCACACCTGCACTCTCGC
PRRX1HD.Q50K.top	GAGTGCAGGTGTGGTTTAAGAACCGAAGAGCC
PRRX1HD.Q50K.bottom	GGCTCTTCGGTTCTTAAACCACACCTGCACTC
PITX2HD.K50S.top	GTCCGGGTTTGGTTCAGCAATCGTCGGGCC
PITX2HD.K50S.bottom	GGCCCGACGATTGCTGAACCAAACCCGGAC
PITX2HD.K50Q.top	GTCCGGGTTTGGTTCCAGAATCGTCGGGCC
PITX2HD.K50Q.bottom	GGCCCGACGATTCTGGAACCAAACCCGGAC

Primers are shown in 5’-3’ direction; bases that introduce a mutation are underlined.

medium containing 5 mg Ampicillin; this culture was grown at 37°C to an OD_{600} of 0.6. Expression of recombinant protein was induced by the addition of 50 µL 100 mM IPTG to the culture followed by a 1 hr incubation at 37°C. After centrifugation for 20 min at 4°C, pellets were resuspended in a mixture of lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, 10 mM imidazole), 100 mM phenylmethanesulphonylfluoride and protease inhibitor cocktail (Sigma). Lysozyme (1 mM) was added to a final concentration of 60 µM and samples were left on ice for 30 min. Whole cell extracts were prepared by sonication (2 x 30 s each with a 10 s cooling period between each pulse). Following centrifugation for 20 min at 4°C, supernatant was transferred to a fresh tube. Recombinant proteins were purified using affinity chromatography on Ni-NTA agarose beads according to the manufacturer's instructions (Qiagen) and concentration of purified protein was determined using the Bradford method.

2.4 Electrophoretic mobility shift assays

Double stranded oligonucleotides used in mobility shift assays (see Table 6) contain a 5'-dG overhang that allows end labeling with α-[^{32}P]-dCTP (3000 Ci/mmol, Amersham) and the Klenow fragment of DNA polymerase (Gibco-BRL). Unincorporated nucleotide was removed using a Nick-Column (Pharmacia-Amersham). Assays were performed in a 20 µL reaction volume containing approximately 4 pmol ^{32}P-labeled oligonucleotide, EMSA buffer (2 mM Tris-HCl (pH 7.5), 10 mM KCl, 1mM DTT, 0.2 mg/mL BSA, and 2% glycerol). One µg p(dI•dC)-p(dI•dC) was added as a nonspecific competitor, as previously described (Underhill and Gros 1997). Protein-DNA complexes were allowed to form at room temperature for 30 min and were resolved

Table 6. Oligonucleotides used for electrophoretic mobility shift assays.

Name	Top strand sequence (5'-3')
PDopt	GGTCGTCACGCTTCAGTGCCCCAT
Nf3'	GCTAGTGTGTGTCACGCTTATTTTCCTGTACTT
P1/2	GATCCTGAGTCTAATTGAGCGTCTGTA
P2	GATCCTGAGTCTAATTGATTACTGTACAG
MITF	GTCATCTTTAGTTCCAGTAGTATTAATAGACAA
Mitf	GTCATCTTTCGTTCCAGTAGTATTAATGGACAA
MITF PDmut	GTCATCTTTACTGATAGTAGTATTAATAGACAA
MITF hd1/2mut	GTCATCTTTAGTTCCAGTACAGTGCCCCATCAA
MITF hd4mut	GGCCGCTTTAGTTCCAGTAGTATTAATAGACAA
MITF hd3mut	GTCATCTGCTGTTCCAGTAGTATTAATAGACAA
MITF hd3top	GTCATCAAATGTTCCAGTAGTATTAATAGACAA
MITF hd3+1	GTCATCTTTAGGTTCCAGTAGTATTAATAGACAA
MITF gaaa	GTCATCTTTCGTTCCAGTACAGTGCCCCATCAA
MITF tgaa	GTCATCTTCAGTTCCAGTACAGTGCCCCATCAA
MITF taga	GTCATCTCTAGTTCCAGTACAGTGCCCCATCAA
MITF taag	GTCATCCTTAGTTCCAGTACAGTGCCCCATCAA
MITF hd3opt	GTCATGATTAGTTCCAGTACAGTGCCCCATCAA
MITF/Nf3'	GTCATCTTTAGTCACGCTTGTATTAATAGACAA
Trp-1 MSEu	GAAGGCCAATGTCACACTTGTATTTTCTGTTAG
Trp-1 MSEi	GCACTAATCCCTTCTCACACCAGTGAATTCTCC
c-MET	GACTCGGTCCCGCTTATCTCCGG

c-RET	GCCAACCACCATGTCACACTGCCCATGGGAGGG
Dct	GTGCACTTAGGGTCATGTGCTAACAAAGAGGAT
MYF5	GTACCATGCAATTAGTCATGCTTTTATGATTTA
Myf5	GTACCATGCAATTAGTCATGCCTTTATGATTTA
SRE	GGACGCAGATGTCCTAATATGGACATCCTGTGT
bicoid	GGATCCGCACGGCCCATCTAATCCCGTGGGATC

All top and bottom strand oligonucleotides have a 5' guanine residue that creates an overhang for ^{32}P-dCTP incorporation.

by electrophoresis (12 V/cm in 0.25X TBE). Resolution of multiple protein-DNA complexes or complexes containing small proteins (e.g., PDs, HDs) was performed with a 10% nondenaturing polyacrylamide gel (29:1 acrylamide:bisacrylamide); resolution of complexes containing large proteins (>20 kDa) was performed with an 8% gel. Polyacrylamide gels were dried under vacuum and exposed overnight (approximately 16 h) to Kodak Biomax MR film with an intensifying screen at -80°C.

2.5 Bulk chromatin isolation

Bulk chromatin from rat liver was obtained using the method of Li et al. (1999b). Briefly, two adult rat livers were minced and homogenized at 4°C. The homogenate was then centrifuged for 30 min at 24,000 rpm at 0°C using a SW 28 rotor (Beckman) to obtain nuclear pellets. Pellets were resuspended and centrifuged at 7000 rpm at 2°C for 4 min to obtain nuclear material. This material was resuspended and pooled, and an aliquot was measured (A_{260} units) to obtain the concentration. The nuclear suspension was diluted to 50 A_{260} units/mL, aliquoted, and stored at -80°C.

2.6 Affinity chromatography pull-down assays

Recombinant, hexahistidine-tagged PAX3-derived proteins (PD, HD, HD+linker, PDHD) or PAX3, PRRX1, or PITX2 HDs were expressed and purified as described (see Section 2.3), but were not eluted. Approximately 0.18 μg of protein-bead slurry was transferred to a new tube, combined with 200 μL micrococcal nuclease (MNase)-digested rat liver chromatin, and incubated for 2 h at 4°C. Beads were washed 4X with Ni-NTA wash buffer (50 mM NaH_2PO_4, 300 mM NaCl, 20 mM imidazole) for 10 min at 4°C.

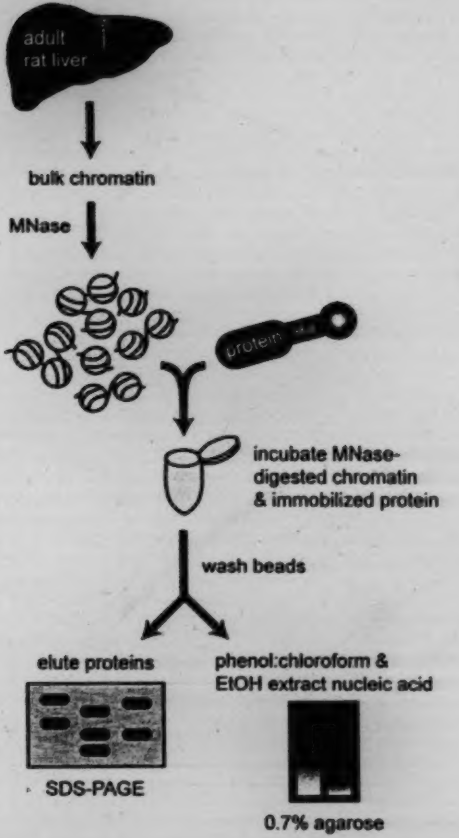

adult
rat liver
bulk chromatin
MNase
protein
incubate MNase-
digested chromatin
& immobilized protein
wash beads
elute proteins
phenol:chloroform &
EtOH extract nucleic acid
SDS-PAGE
0.7% agarose

Figure 2-1. Affinity chromatography pull-down assays and analysis of PAX3- or Paired-type homeodomain interaction with chromatin and associated nucleic acid. Bulk chromatin from rat liver was used to investigate chromatin interaction with immobilized recombinant PAX3-derived proteins or Prd-type HDs. Chromatin was digested with micrococcal nuclease (MNase), combined with hexahistidine-tagged protein conjugated to nickel-agarose beads, and washed to dissociate weakly interacting proteins. Beads were either eluted in SDS buffer and resolved using polyacrylamide gel electrophoresis (PAGE), or protein was phenol:chloroform extracted and nucleic acid ethanol precipitated, then resolved on a 0.7% agarose gel. See Sections 2.5, 2.6, and 5.4 for details.

Proteins were eluted in SDS sample buffer, resolved by SDS-PAGE, and stained with Coomassie blue. Interacting chromatin-associated DNA was analyzed by phenol:chloroform extraction, ethanol precipitation, and resolution on a 0.7% agarose gel (summarized in Figure 2-1).

2.7 Cell lines and culture

Mouse embryonic fibroblast C3H 10T1/2 cells and B16F10 murine melanoma cells were obtained from American Type Culture Collection. HeLa cells were kindly provided by the laboratory of Dr. M. Walter, University of Alberta. Cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 2 mM L-glutamine and 10% fetal bovine serum at 37°C and 5% CO_2.

2.8 Dual luciferase assays

HeLa cells were plated in 24-well dishes and grown for 24 h to approximately 70% confluency. Cells were transiently transfected using FuGENE 6 (Roche) with PAX3 expression plasmids, the Renilla luciferase pRL-TK control plasmid (Promega), and the pGL3-basic plasmid or the *MITF-Luc* or *Trp1-Luc* reporter constructs (see Section 2.1.3). Luciferase activity was quantified 24 h post-transfection using the Dual-Luciferase Reporter Assay System^{TM} (Promega) in a TD 20/20 luminometer (Turner Designs). Statistical comparisons were made using the Student's paired t-test with a two-tailed distribution. A p value of <0.05 was considered statistically significant.

2.9 Antibodies

A PAX3-specific monoclonal antibody developed by C. Ordahl (University of California, San Francisco) and obtained from the Developmental Studies Hybridoma Bank (DSHB), NICHD/University of Iowa, was used for western analysis (1:1500) and immunofluorescence (1:200). Other antibodies used for immunofluorescence were specific for bromodeoxyuridine (BrdU) (1:1000; Sigma), acetylated histone H3 (AcH3) (1:200; Upstate), acetylated histone H4 (AcH4) (1:200; Upstate), trimethyl-lysine 4 of histone H3 (H3K4me3) (1:500; Abcam), trimethyl-lysine 36 of histone H3 (H3K36me3) (1:200; Abcam), trimethyl-lysine 9 of histone H3 (H3K9me3) (1:500; Upstate), trimethyl-lysine 20 of histone H4 (H4K20me3) (1:200; Abcam), and TLE4/GRG4 (1:400, Sigma).

2.10 Whole cell extracts

Murine B16F10 melanoma cells and 10T1/2 fibroblasts were used for whole cell extracts. Cells were seeded on a 15-cm plate and grown to approximately 80% confluency. Transfections in 10T1/2 cells were performed with 10 µg pcDNA3.1-PAX3 in DMEM containing 50 µL polyethylenimine (PEI) or left untransfected. B16F10 cells were not transfected. After 24 h, cells were washed 2X with ice cold PBS, collected, and lysed in lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% IGEPAL). Lysates were sonicated for 2 x 12 s at low power, and then centrifuged at 13,000 rpm at 4°C for 10 min. Supernatant was aliquoted and stored at -80°C. Western analysis (see Section 2.11) was used to confirm PAX3 expression.

2.11 Western transfer

Whole cell extracts (see Section 2.10) were boiled in SDS sample buffer with β-mercaptoethanol then resolved by SDS-PAGE. Proteins were blotted onto Trans-Blot Transfer Medium (Bio-Rad) and blocked for 1 h at room temperature in TBST (25 mM Tris-HCl (pH 8.0), 0.15 M NaCl, 0.05% Tween) containing 5% skim milk. Blots were incubated with the PAX3-specific DSHB primary antibody diluted 1:1500 in TBST/milk for 1 h at room temperature, then washed 4X with TBST. Following this, blots were incubated with a secondary antibody conjugated to peroxidase for 1 h, and washed again with TBST/milk. Blots were washed 2X with TBS, treated with enhanced chemiluminescence reagent (Amersham), and exposed to Kodak Biomax MR film (Eastman Kodak Co.) for 10 min.

2.12 Immunocytochemistry

Cells were seeded on glass cover slips in 6-well plates at a density of ~1.0×10^6 cells/ml. After 24 h growth, cells were transfected with 1 μg DNA in DMEM containing FuGENE 6 (Roche) or PEI. At 24 h post-transfection, cells were washed in PBS and fixed for 5 min at room temperature with 4% paraformaldehyde in PBS. Cells were then permeabilized with 10% Triton X-100 in PBS. Cells transfected with GFP constructs were mounted onto glass slides with 20 μL Mowiol containing 1μg/ml Hoechst 33258 (Sigma). For immunofluorescence, cells were blocked using 5% PBS-BSA for 1 h and washed 3X with 1% PBS-BSA prior to application of primary antibodies (see Section 2.9) diluted in 1% PBS-BSA. Cells were washed 3X with 1% PBS-BSA then incubated with a secondary antibody (Sigma; 1:400 in 1% PBS-BSA) conjugated to a fluorescent moiety for 1 h. Coverslips were washed with PBS and mounted onto glass slides with 20 µL Mowiol containing Hoechst. Fluorouridine (5-FUrd) staining of B16F10 cells or PAX3-transfected 10T1/2 cells was performed by diluting 5-FUrd in DMEM to a final concentration of 1 mM, then adding this mixture to cells and incubating for 30 min at 37°C.

Epifluorescence imaging was performed using a Zeiss Axioplan 2 optical microscope equipped with a Photometrics CoolSnap HQ CCD camera (Roper Scientific Inc.). Images were captured using MetaMorph 2.6r6 (Molecular Devices) and processed with Huygens deconvolution software (Scientific Volume Imaging). Three-dimensional rendering and image construction was performed using Imaris (Bitplane AG).

2.13 Fluorescence recovery after photobleaching

PAX3-GFP constructs were transfected into 10T1/2 cells and images were collected with a Zeiss Laser Scanning Confocal Microscope (LSM 510, software version LSM 3.2) mounted on a Zeiss Axiovert M100 inverted microscope with a 40X apochromatic lens (numerical aperture 1.3). The 488 nm laser line (from a 25 mW argon laser) was used to image the PAX3-GFP constructs. A long pass filter (505 nm) was used to collect emission from the PAX3-GFP constructs. A 2 µM rectangle was bleached across the center of the nucleus and images were taken at 1 s intervals for the first 4 s, 2 s intervals from seconds 4 to 14, and 5 s intervals from seconds 14 to 90 following photobleaching.

2.14 Bioinformatics

Evolutionary analysis of PAX3 target promoter sequences was performed by comparing sequences containing the putative PAX3-recognition elements of the human *MITF*, *MYF5*, and *c-MET*, and mouse *Dct* promoters to the Whole-genome shotgun reads (wgs) database in the BLAST program (Altschul et al. 1990). Sequences were aligned with ClustalW (http://www.ebi.ac.uk/clustalw/) to obtain homology and shaded using BoxShade 3.21 (http://www.ch.embnet.org/). Graphical analysis of base preference within sequences was obtained with WebLogo 2.8.2 (http://weblogo.berkeley.edu/).

The distribution of 5'-TAATNN sites was examined using mouse genome sequence based on NCBI Build 36 (www.ensembl.org). Chromosome-specific FASTA sequence files (unmasked and repeat masked) were analyzed using a PERL program that searched for all sixteen iterations of TAATNN on both DNA strands. In addition, genomic sequences were characterized using a PERL program that monitored the positions of TAATNN motifs, which could then be used to determine motif density along the chromosome. For both programs, data files were logged to Microsoft Excel for analysis and comparison to Ensembl genome annotations (www.ensembl.org).

CHAPTER 3. PAX3-TARGET INTERACTION

Portions of this chapter are published in:

Corry GN, Underhill DA. (2005) Pax3 target gene recognition occurs through distinct modes that are differentially affected by disease-associated mutations. *Pigment Cell Res.* **18**: 427-38.

3.1 Background

PAX3 controls the expression of downstream target genes by interacting with regulatory elements in these genes' promoters. The discovery of an optimized binding site for the PAX3 PD (*Nf3'*; Epstein et al. 1995) has led to the identification of several putative regulatory sites in genes that are genetically downstream of PAX3. These include the tyrosine kinase receptors *c-MET* (Epstein et al. 1996) and *c-RET* (Lang et al. 2000), the *Msx2* (Kwang et al. 2002), *Myf5* (Bajard et al. 2006), and *MITF* (Watanabe et al. 1998) transcription factors, and the melanin synthetic enzyme-encoding *Trp-1* (Galibert et al. 1999) and *Dct* (Lang et al. 2005) genes (see Section 1.1.7). Based on genetic and biochemical studies, these genes represent suitable candidates for PAX3 target genes but so far, only *Myf5* has been shown to be directly bound and regulated by PAX3 *in vivo*. Furthermore, any inappropriate effects on these genes' regulation in the context of mutant or null alleles of PAX3 are poorly understood. The majority of *PAX3* mutations occur within either the PD or HD (see Fig. 1-4), suggesting that the primary defect in these cases relates to DNA binding (Underhill 2000). Thus, it is important to clarify not only how PAX3 regulates expression of target genes, but also how PAX3 mutations affect this process in the interest of understanding how PAX3-associated disorders arise.

In this chapter, we have performed a detailed analysis of how PAX3 interacts with characterized promoter sequences. We find that PAX3 targets can be classified depending on the presence of a single PD binding site or a composite sequence containing recognition motifs for both the PD and HD. Affinity of PAX3 for putative *in vivo* targets varies and, while the PD is necessary for PAX3 DNA binding, the HD makes an important contribution to high affinity interaction with composite elements. Surprisingly, a consensus recognition motif is not necessary for the HD to interact with composite sequences. We also assess the effects of disease mutations on PAX3 DNA binding and establish that they exert distinct effects on PAX3-DNA recognition that support mechanistically distinct interactions with the DNA. Importantly, these mutations also exert distinct effects on the ability of PAX3 to regulate reporter genes fused to either the *MITF* or *Trp-1* promoters. Thus, our results establish that PAX3 can regulate target genes through alternate modes of DNA recognition that are differentially impacted by disease-causing mutations, which together have important implications for understanding PAX3-regulated gene networks.

3.2 Analysis of PAX3 recognition sequences

The characterization of putative PAX3 targets has revealed two apparent types of recognition sequences (Fig. 3-1). One type, which resembles the *in vitro*-optimized *Nf3'* site (Epstein et al. 1995), as well as elements from the sea urchin *H2A-2/H2B-2* genes (Barberis et al. 1989) and the B-cell specific *CD19* gene (Kozmik et al. 1992), features a sequence with a 5'-GTCAC-like PD-binding motif. Analogous PAX3 targets include *c-MET*, *c-RET*, *Trp-1*, *Msx2*, and *Dct*. The second type more closely resembles the *e5* sequence (Hoey and Levine 1988) and derivative *paired domain recognition sites* (*PRS*) (Chalepakis et al. 1991), which feature a PD-binding site and a putative HD-recognition motif; targets of this class include the *MITF*, *MYF5*, and *Ng-CAM* genes. To date, no functional target sequences have been identified that contain recognition sites for the PAX3 HD only, although mutation of the TAAT motif in the *Ng-CAM* promoter element

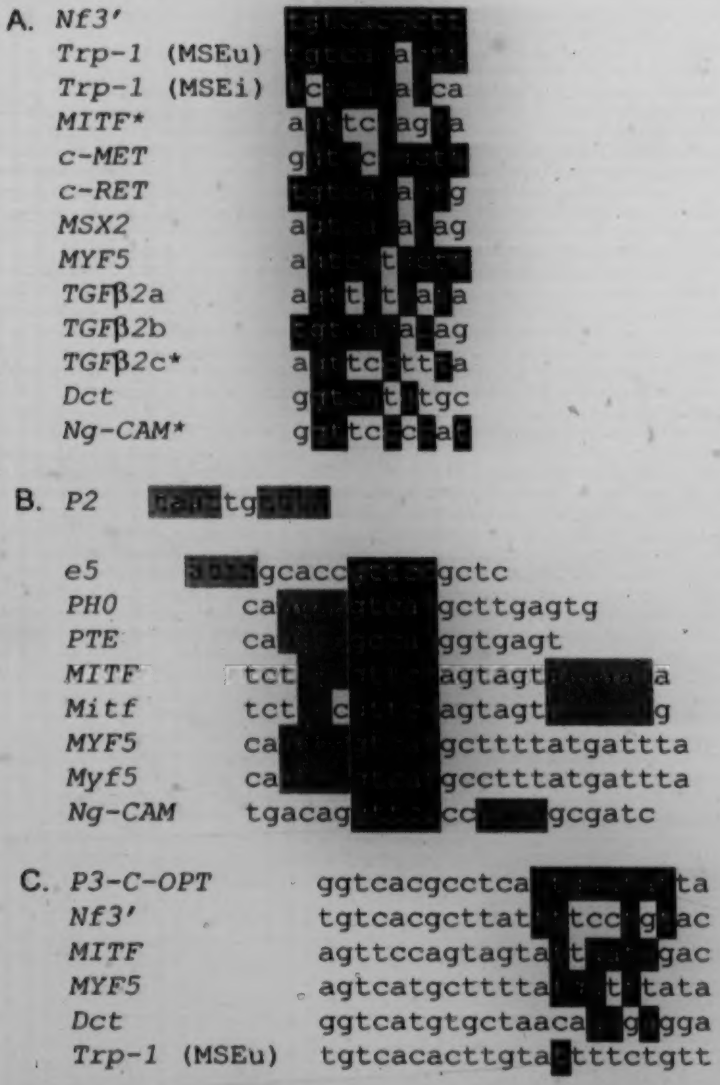

A.
Nf3'
Trp-1 (MSEu)
Trp-1 (MSEi)
MITF*
c-MET
c-RET
MSX2
MYF5
TGFβ2a
TGFβ2b
TGFβ2c*
Dct
Ng-CAM*
B.
P2
e5
PHO
PTE
MITF
Mitf
MYF5
Myf5
Ng-CAM
C.
P3-C-OPT
Nf3'
MITF
MYF5
Dct
Trp-1 (MSEu)

Figure 3-1. Alignment of selected PAX3 target recognition elements. (A) Alignment of PD-binding motifs from putative PAX3 target promoter sequences with the *in vitro*-optimized PAX3 PD binding site *Nf3'* (*top*). Bases conserved between *Nf3'* and *in vivo* sites are highlighted in black. Asterisks denote GTTCC-type motifs similar to that of the *e5* sequence. (B) Alignment of composite target sequences that feature potential recognition motifs for both the PAX3 PD and HD. The *in vitro*-optimized Prd-class HD binding site *P2* is shown at the top. Putative PD-binding sites are highlighted in black; putative HD recognition sites are highlighted in grey. (C) Alignment of selected *in vivo* PAX3 targets with the optimized carboxy-terminal paired subdomain recognition sequence (*P3-C-OPT*). Bases sharing identity across the carboxy-terminal subdomain binding region are highlighted in black.

abrogated binding by PAX3 (Kallunki et al. 1995). Together, these data suggest that PAX3 regulates the expression of a diverse set of target genes using two distinct DNA binding modes – one based on PD-only contact and the other involving the PD and HD.

Alignment of putative *in vivo* PAX3 targets with the *Nf3'* sequence reveals a striking variation in the degree of conservation between PD recognition motifs (Fig. 3-1A). *Trp-1* (MSEu), *c-RET*, *MSX2*, and *TGFβ2b* are the only sequences that feature the GTCAC core found in the *Nf3'* sequence. The *MYF5*, *c-MET*, and *Dct* sites resemble the *Nf3'* sequence, but contain substitutions at key positions within the GTCAC motif. The *MITF*, *TGFβ2*c, and *Ng-CAM* sites possess a GTTCC motif, similar to that found in the *e5* sequence, but none of these sequences contain the 5'-ATTA HD-binding element characteristic of *e5* (see below). Thus, only a few suspected targets of PAX3 feature an optimal PD-binding motif, suggesting that either the PAX3 PD is remarkably flexible in its ability to interact with various recognition sequences or that a canonical PD-recognition site is dispensable for optimal binding of PAX3 to target promoters. It must be considered, however, that most analyses of PAX3 binding to target sequences have been performed under *in vitro* conditions and that affinity *in vivo* is likely to be influenced by chromatin structure and other nuclear components.

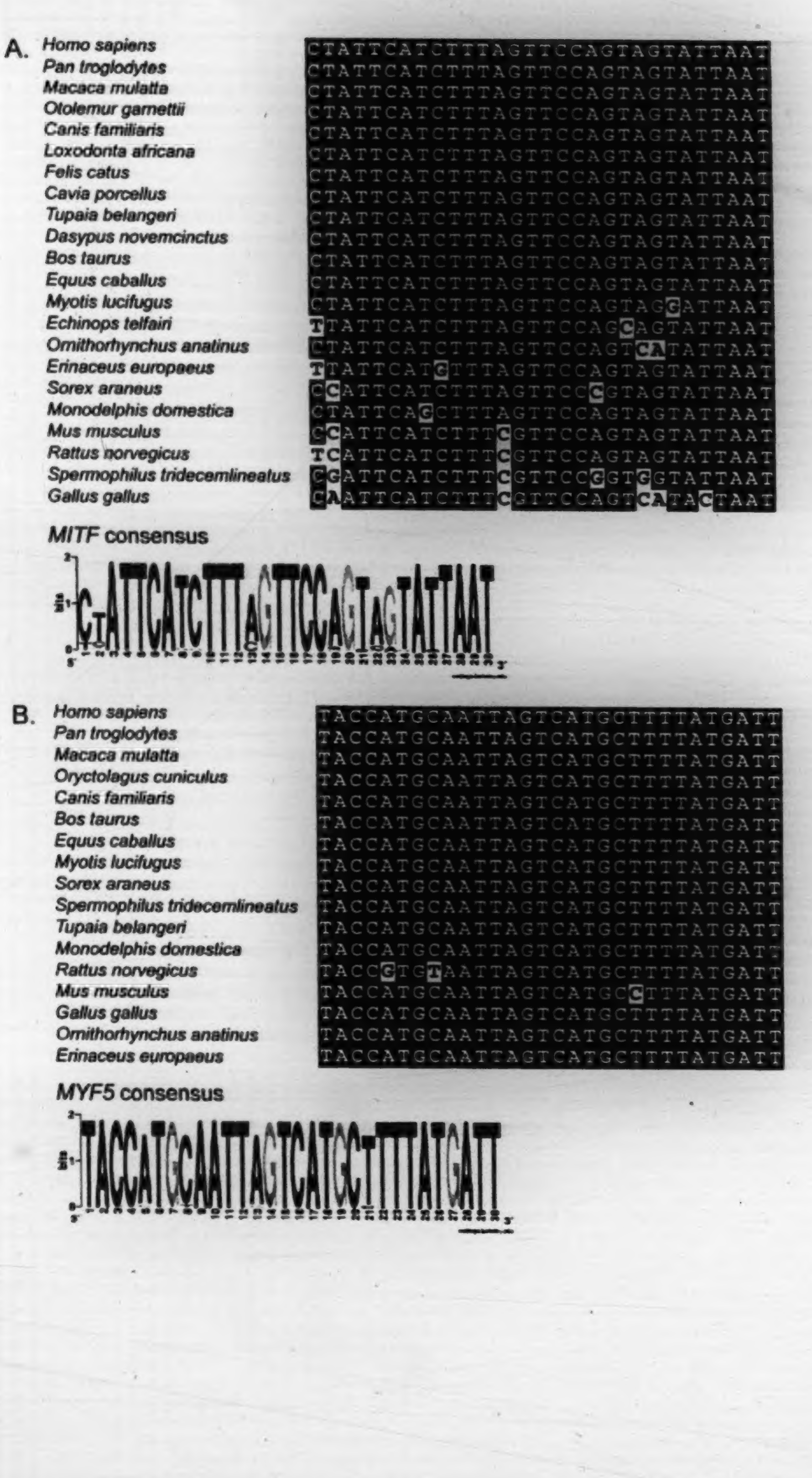
The PAX3 binding elements of the *MITF*/*Mitf* and *MYF5*/*Myf5* promoters possess both a PD recognition motif and one or more putative HD-binding sites (Watanabe et al. 1998; Bajard et al. 2006). We therefore undertook a closer examination of these sites in comparison to *e5* and composite PD-HD sequences described by others (Fig. 3-1B). The *e5* sequence (Section 1.1.3) and *PTE*/*PH0* sites, which confer cooperative binding by the Prd PD and HD (Section 1.1.9), have been described previously. The *Ng-CAM* promoter site contains putative PD and HD recognition motifs and is therefore included as a composite sequence, but only the HD site appeared to be involved in binding PAX3 (Kallunki et al. 1995). Alignment of these sequences reveals an interesting trend: with the exception of the *e5* and *Ng-CAM* sequences, a completely or partially conserved HD-recognition motif is immediately adjacent to the PD motif on the 5' side. In the case of *e5*, the 5 bp spacing between the PD and HD motifs would separate them by half a turn of the DNA helix, possibly inducing a cooperative binding event that maintains high affinity binding of both domains, albeit in an alternate nature to the sites that contain juxtaposed PD and HD motifs. There is also no conservation in the PD recognition motif among the composite sequences, as both GTCAC- and GTTCC-like motifs are found.

The *MITF/Mitf* and *Myf5/MYF5* elements each contain AT-rich sequences downstream of the PD site and, in the case of the former, feature consensus HD recognition motifs. However, these regions also bear resemblance to the *in vitro*-optimized sequence for the PAX3 carboxy-terminal PD subdomain (*P3-C-OPT*), as determined by (Vogan and Gros 1997) (Fig. 3-1C). Both the *MITF* and *MYF5* elements share considerable identity across the core carboxy-terminal subdomain-binding region, but the *Dct* and *Trp-1* MSEu sites, both PD-only motifs, feature less similarity with *P3-C-OPT* within this region (Fig. 3-1C, *highlighted sequence*). Together, these observations provide compelling evidence for a distinct set of PAX3 target sequences that require both PD and HD interaction and suggest that the amino- and carboxy-terminal PAX3 PD subdomains may be employed in different manners, depending on the recognition sequence.

3.3 Evolutionary conservation of selected PAX3 target sequences

We next performed *in silico* phylogenetic analysis of selected *in vivo* PAX3 binding sites to determine the extent of sequence conservation across evolution (Fig. 3-2). Two composite sequences, *MITF* and *MYF5*, and two PD-only sequences, *Dct* and *c-MET*, were examined. Fig. 3-2A shows that the PAX3 recognition element of the *MITF* promoter is highly conserved among higher eukaryotes. In particular, the GTTCC PD motif is present in each sequence, as is a triplet of thymine residues one bp upstream of the PD motif. In rodents, a cytosine separates these elements, while in higher organisms, an adenine is found. As will be shown later, neither base confers a difference in PAX3 affinity, suggesting this position is inconsequential for recognition of the *MITF* promoter sequence. The PAX3 binding site of the *MYF5* promoter is similarly highly conserved among higher eukaryotes (Fig. 3-2B). A perfectly conserved GTCAT PD motif and an adjacent consensus HD-binding motif (ATTA) are present in all organisms that were assessed. Also conserved in the *MITF* and *MYF5* sequences is the extended AT-rich region that follows the PD-binding core motif and is thought to bind the carboxy-terminal PD subdomain (see Fig. 3-1C).

The proposed PAX3-binding element of the *Dct* promoter overlaps with an M-box motif that binds MITF (Lang et al. 2005). In general, the 5'-GTCAT PD motif is well conserved and of all the organisms compared, only the *O. cuniculus* and *O. princepus* promoters did not contain a GTCAT motif, although the flanking sequence showed no divergence from the *Dct* consensus in these animals (Fig. 3-2C). In contrast to *Dct*, the *c-MET* promoter region that encompasses the MET1 binding site (Epstein et al. 1996; see Section 1.1.7) is poorly conserved among higher eukaryotes (Fig. 3-2D),


A.
Homo sapiens CTATTCATCTTTAGTTCCAGTAGTATTAAT
Pan troglodytes CTATTCATCTTTAGTTCCAGTAGTATTAAT
Macaca mulatta CTATTCATCTTTAGTTCCAGTAGTATTAAT
Otolemur garnettii CTATTCATCTTTAGTTCCAGTAGTATTAAT
Canis familiaris CTATTCATCTTTAGTTCCAGTAGTATTAAT
Loxodonta africana CTATTCATCTTTAGTTCCAGTAGTATTAAT
Felis catus CTATTCATCTTTAGTTCCAGTAGTATTAAT
Cavia porcellus CTATTCATCTTTAGTTCCAGTAGTATTAAT
Tupaia belangeri CTATTCATCTTTAGTTCCAGTAGTATTAAT
Dasypus novemcinctus CTATTCATCTTTAGTTCCAGTAGTATTAAT
Bos taurus CTATTCATCTTTAGTTCCAGTAGTATTAAT
Equus caballus CTATTCATCTTTAGTTCCAGTAGTATTAAT
Myotis lucifugus CTATTCATCTTTAGTTCCAGTAGGATTAAT
Echinops telfairi TTATTCATCTTTAGTTCCAGCAGTATTAAT
Ornithorhynchus anatinus CTATTCATCTTTAGTTCCAGTCATATTAAT
Erinaceus europaeus TTATTCATGTTTAGTTCCAGTAGTATTAAT
Sorex araneus CCATTCATCTTTAGTTCCCGTAGTATTAAT
Monodelphis domestica CTATTCAGCTTTAGTTCCAGTAGTATTAAT
Mus musculus CCATTCATCTTTCGTTCCAGTAGTATTAAT
Rattus norvegicus TCATTCATCTTTCGTTCCAGTAGTATTAAT
Spermophilus tridecemlineatus CGATTCATCTTTCGTTCCGGTGGTATTAAT
Gallus gallus CAATTCATCTTTCGTTCCAGTCATACTAAT
MITF consensus
CTATTCATCTTTAGTTCCAGTAGTATTAAT
5'
3'
B.
Homo sapiens TACCATGCAATTAGTCATGCTTTTATGATT
Pan troglodytes TACCATGCAATTAGTCATGCTTTTATGATT
Macaca mulatta TACCATGCAATTAGTCATGCTTTTATGATT
Oryctolagus cuniculus TACCATGCAATTAGTCATGCTTTTATGATT
Canis familiaris TACCATGCAATTAGTCATGCTTTTATGATT
Bos taurus TACCATGCAATTAGTCATGCTTTTATGATT
Equus caballus TACCATGCAATTAGTCATGCTTTTATGATT
Myotis lucifugus TACCATGCAATTAGTCATGCTTTTATGATT
Sorex araneus TACCATGCAATTAGTCATGCTTTTATGATT
Spermophilus tridecemlineatus TACCATGCAATTAGTCATGCTTTTATGATT
Tupaia belangeri TACCATGCAATTAGTCATGCTTTTATGATT
Monodelphis domestica TACCATGCAATTAGTCATGCTTTTATGATT
Rattus norvegicus TACCGTGTAATTAGTCATGCTTTTATGATT
Mus musculus TACCATGCAATTAGTCATGCCTTTATGATT
Gallus gallus TACCATGCAATTAGTCATGCTTTTATGATT
Ornithorhynchus anatinus TACCATGCAATTAGTCATGCTTTTATGATT
Erinaceus europaeus TACCATGCAATTAGTCATGCTTTTATGATT
MYF5 consensus
TACCATGCAATTAGTCATGCTTTTATGATT
5'
3'

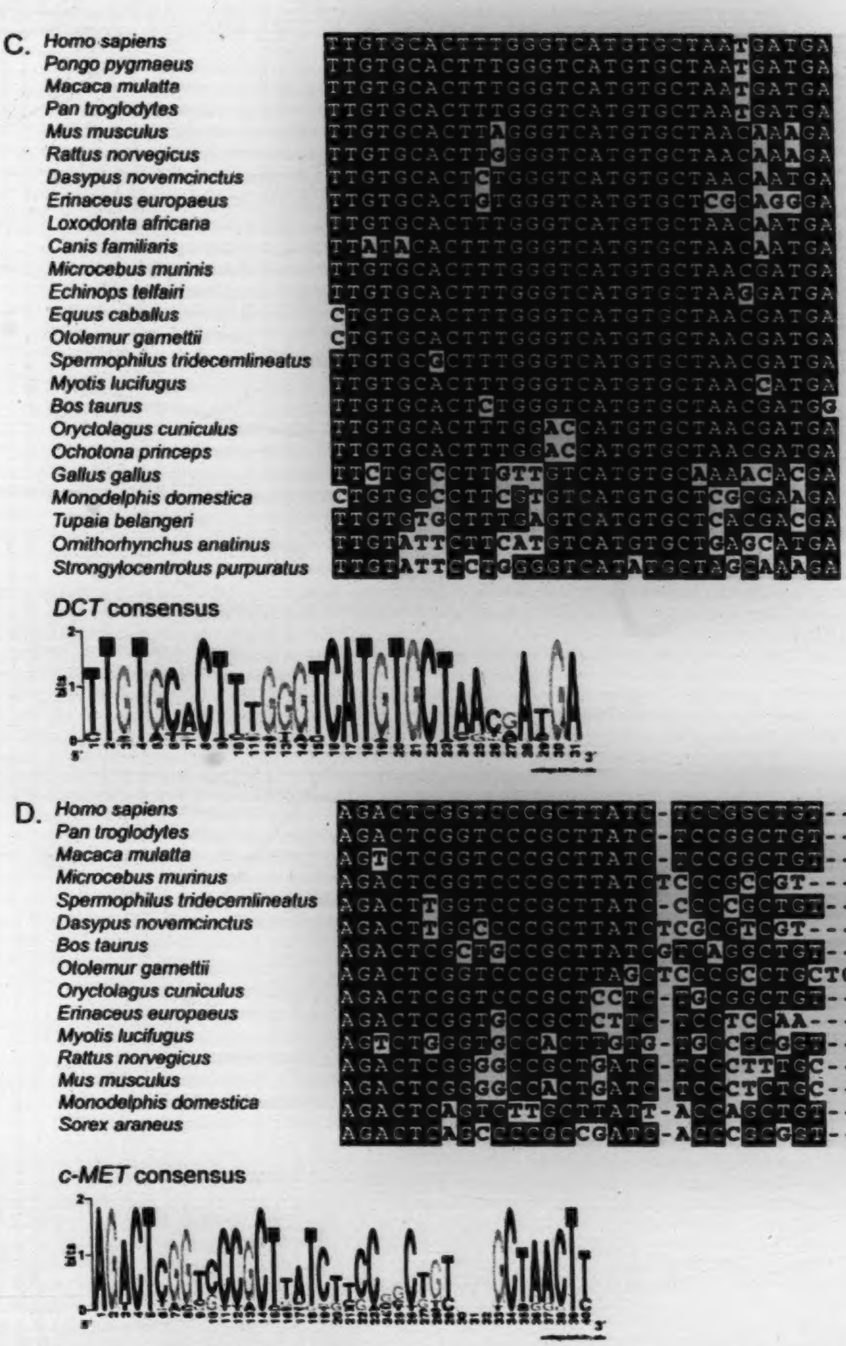

C.
Homo sapiens
Pongo pygmaeus
Macaca mulatta
Pan troglodytes
Mus musculus
Rattus norvegicus
Dasypus novemcinctus
Erinaceus europaeus
Loxodonta africana
Canis familiaris
Microcebus murinis
Echinops telfairi
Equus caballus
Otolemur garnettii
Spermophilus tridecemlineatus
Myotis lucifugus
Bos taurus
Oryctolagus cuniculus
Ochotona princeps
Gallus gallus
Monodelphis domestica
Tupaia belangeri
Ornithorhynchus anatinus
Strongylocentrotus purpuratus
DCT consensus
D.
Homo sapiens
Pan troglodytes
Macaca mulatta
Microcebus murinus
Spermophilus tridecemlineatus
Dasypus novemcinctus
Bos taurus
Otolemur garnettii
Oryctolagus cuniculus
Erinaceus europaeus
Myotis lucifugus
Rattus norvegicus
Mus musculus
Monodelphis domestica
Sorex araneus
c-MET consensus

Figure 3-2. Evolutionary conservation of PAX3 target promoter sequences. *In silico* analysis of the proposed PAX3-binding elements of the *MITF* (A), *MYF5* (B), *Dct* (C), and *c-MET* (D) promoters was performed to compare sequence conservation among higher eukaryotes. Bases that share identity are highlighted in black. Sequences were aligned using ClustalW and shaded with BoxShade 3.21. Relative conservation at each position within the sequences was obtained using WebLogo 2.8.2 (*bottom*).

including the putative 5'-GTCCC PD motif, suggesting that the MET1 site is not a functional PAX3-responsive element. Interestingly, the *Trp-1* promoter region containing the MSEi and MSEu motifs showed no evolutionary conservation (not shown), despite, as will be demonstrated below, its ability to bind PAX3.

3.4 DNA binding characteristics of PAX3

To examine the binding requirements of PAX3 to these diverse recognition motifs, we performed mobility shift assays using COS-7 whole cell extracts containing the full length PAX3 Q+ and Q- isoforms. COS-expressed PAX3 has previously been shown to possess typical DNA binding activity (Underhill et al. 1995; Vogan et al. 1996). We analyzed binding of the Q+ and Q- isoforms (see Section 1.1.2) to oligonucleotides containing the PAX3-binding elements of the *MITF*, *Trp-1* (MSEu), and *c-MET* promoters, as well as an oligonucleotide containing an optimized PD-binding site (*PDopt*) (Fig. 3-3A; see Table 6 for oligonucleotide sequences). Both isoforms bound with high affinity to the *PDopt* and *MITF* oligonucleotides and with slightly lower affinity to the *Trp-1* sequence. Neither isoform bound to the *c-MET* oligonucleotide. These results confirm that full length PAX3 binds with comparable affinity to both GTCAC-type (*Trp-1*) and GTTCC-type (*MITF*) PD motifs, as previously demonstrated by Vogan et al. (1996). The same group showed that the Q+ and Q- isoforms exhibited differing affinities depending on the presence (*P6CON* and *CD19-2/A*) or absence (*e5*) of a recognition sequence for the carboxy-terminal PD subdomain (Vogan et al. 1996). Our data show a similar difference between Q+ and Q- binding to the *MITF* and *Trp-1* oligonucleotides, suggesting that these sequences also rely on a contribution from the

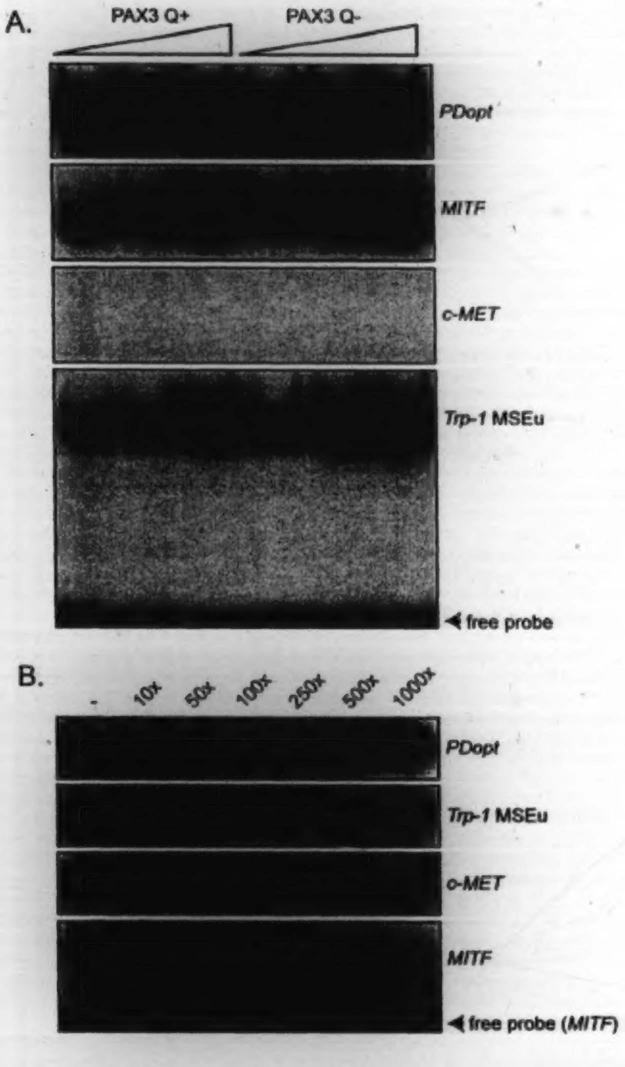

A.
PAX3 Q+
PAX3 Q-
PDopt
MITF
c-MET
Trp-1 MSEu
free probe
B.
-
10x
50x
100x
250x
500x
1000x
PDopt
Trp-1 MSEu
c-MET
MITF
free probe (MITF)

Figure 3-3. Interaction of PAX3 with selected target promoter sequences. (A) Whole cell extracts from COS-7 cells transiently transfected with PAX3 Q+ or PAX3 Q- expression constructs were used in mobility shift assays with ^{32}P-labeled oligonucleotides containing an optimized PAX3 PD-binding site (*PDopt*) or the PAX3-binding elements from the *MITF* and *c-MET* promoters, as well the *Trp-1* promoter MSEu site (see Table 6 for oligonucleotide sequences). Increasing amounts of extract used in each lane are as indicated; free probe is shown for the bottom panel only. (B) Relative affinity of PAX3 for *in vitro*- (*Nf3'*) and *in vivo*-derived (*Trp-1* MSEu, *c-MET*, *MITF*) target sites was assayed by competition mobility shift assays using the ^{32}P-labeled *MITF* oligonucleotide in the presence of increasing concentrations of the indicated competitor oligonucleotide (free probe is indicated for the bottom panel only). Competitor DNA was added in 10, 50, 100, 250, 500, or 1000-fold molar excess, as indicated above each lane. The PAX3 PDHD protein was used at a fixed concentration in all panels shown.

carboxy-terminal subdomain. Lastly, our data demonstrate that deviation at a heavily selected position within the core PD motif (i.e., *c-MET*) is sufficient to abrogate binding. A PD-binding site selection assay revealed that the adenine of 5'-GTC$\underline{A}$C is strongly preferred by the PAX3 PD (Epstein et al. 1996), likely accounting for the failure of PAX3 to bind to the *c-MET* oligonucleotide.

We next compared the affinity of PAX3 for the same binding sequences and established that a minimal recombinant protein containing the PAX3 DNA-binding domains (PDHD) can efficiently recapitulate the DNA binding properties of the full length protein. We analyzed the ability of unlabeled oligonucleotides containing the four sequences used in Fig. 3-3A to compete for PAX3 in the presence of a radiolabeled *MITF* oligonucleotide. We chose to label the *MITF* oligonucleotide since it mediated strong PAX3-binding activity (Fig. 3-3A). For these assays, we used a recombinant protein containing the PD and HD tethered by the native interdomain linker (PDHD; Fig. 3-4A below). Of these targets, the PDHD protein displayed the highest affinity for the *MITF* sequence (Fig. 3-3B). Surprisingly, the PDHD protein showed less affinity for the *PDopt* oligonucleotide compared to *MITF*, possibly due to the contribution of HD binding to the latter sequence. This result also suggests that regions outside of that encompassed by the PDHD protein contribute to PAX3 DNA binding, since the full length protein showed higher affinity for *PDopt* than *MITF* (see above). Similar to *PDopt*, the *Trp-1* MSEu oligonucleotide did not compete for PDHD as efficiently as *MITF*, but we did observe a moderate level of competition at high *Trp-1* concentrations. Consistent with experiments using the full length protein, we did not observe any competition by the *c-MET* oligonucleotide. Thus, PAX3 appears to bind with higher affinity to a composite PD-HD sequence (*MITF*) than to sequences that contain binding sites for the PD alone.

3.5 The PAX3 homeodomain increases binding affinity for composite sequences

Analysis of currently proposed PAX3-binding sequences suggests the existence of two distinct types: those that feature motifs for binding by the PD, and those that contain both PD- and HD-binding sites. The functional relevance of most of these sequences *in vivo* remains poorly characterized; however, the existence of two distinct groups of recognition motifs presents a possible method of differentially regulating PAX3 target gene expression by conferring distinct binding modes. Above, we showed that PAX3 binds efficiently to the *MITF* and *Trp-1* MSEu sites, despite their differences in sequence composition (and in spite of the lack of conservation of the *Trp-1* sequence). To gain a better understanding of how PAX3 interacts with PD-only (*Trp-1* MSEu, *c-RET*, *Dct*) and composite (*MITF*, *MYF5*) promoter sequences, we performed mobility shift assays with additional representatives of each group. We used recombinant proteins encompassing the PD or PDHD (see Fig. 3-4A) to clarify how PAX3 makes use of the PD and HD, particularly in the presence of composite motifs. The PD bound poorly to all sites except for *MYF5* (Fig. 3-4B, *top panel*), despite the *Trp-1* and *c-RET* sequences containing consensus 5'-GTCAC motifs. At an equal concentration to that used for the PD, the PDHD showed improved binding to *Trp-1*, *MITF*, and *MYF5*, but possessed low (*c-RET*) or no (*Dct*) affinity for the remaining sequences (Fig. 3-4B, *bottom panel*). These results suggest that tethering the HD to the PD improves the affinity of PAX3 for composite sites, but has negligible effects on binding to PD-only sequences.

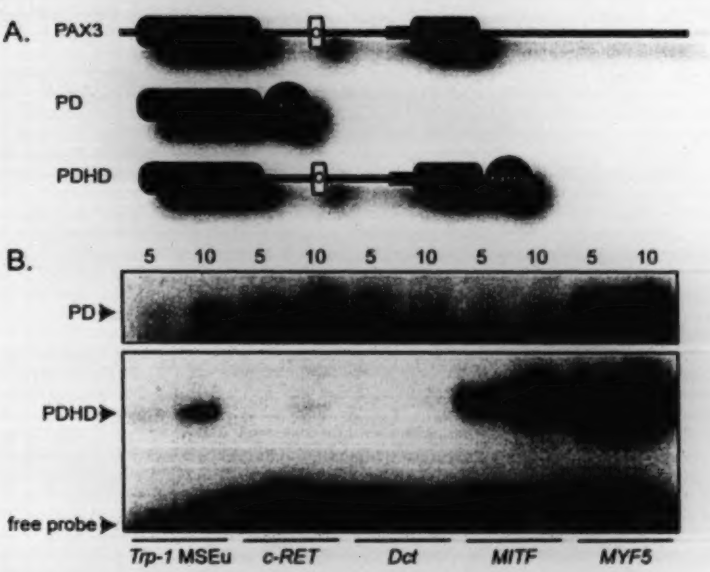

A.
PAX3
PD
PDHD
B.
5
10
5
10
5
10
5
10
5
10
PD
PDHD
free probe
Trp-1 MSEu
c-RET
Dct
MITF
MYF5

Figure 3-4. The homeodomain contributes to PAX3 interaction with composite target elements. (A) Schematic of recombinant PAX3-derived proteins used in mobility shift assays to investigate structural determinants of PAX3 DNA binding. PD, paired domain; HD, homeodomain; O, octapeptide; NLS, nuclear localization signal; 6xHis, hexahistidine tag. See Section 2.1.1 for details. (B) PAX3 PD (*top*) and PDHD (*bottom*) binding to selected PD-only (*Trp-1*, *c-RET*, *Dct*) and composite PD-HD (*MITF*, *MYF5*) sequences. Positions of protein-DNA complexes and free probe are indicated; amounts (µL) of protein added are shown above each lane; all proteins were equalized to 9 ng/µL.

3.6 Analysis of PAX3 interaction with the *Trp-1* promoter sequence

Both the *Trp-1* MSEu and MSEi motifs were shown to be necessary for activation of a luciferase reporter construct containing the *Trp-1* upstream regulatory region (Galibert et al. 1999). However, the roles of these motifs in regulating *Trp-1* expression *in vivo* remain unclear, and the lack of conservation in this part of the *Trp-1* promoter (see Section 3.3) suggests that this region may not contain any functional regulatory elements. Above, we showed that PAX3 binds to the *Trp-1* MSEu site despite its lack of conservation, and that tethering the HD to the PD caused moderate improvement in affinity. In contrast to the MSEu site, the MSEi site contains a 5'-TAAT HD target motif downstream of a non-canonical PD motif, and luciferase reporter assays demonstrated the that the TAAT motif was required for activation of a *Trp-1* promoter-linked gene (Galibert et al. 1999). To clarify PAX3-*Trp-1* interaction, we performed DNA binding assays using oligonucleotides containing either the MSEu or MSEi site (Fig. 3-5A). The PDHD protein bound to both elements with comparable affinity (Fig. 3-5B, *top panel*) but only the MSEu site was bound by the PD (Fig. 3-5B, *bottom panel*), indicating that the HD is entirely responsible for interaction with the MSEi site, as suggested by the results of Galibert et al. (1999). Together with our analysis of the evolutionary conservation of the *Trp-1* promoter, these data support our prediction that neither the MSEu nor the MSEi motif is a functional PAX3 recognition element, despite the ability of the former to bind PAX3 *in vitro*. Given the involvement of both TRP-1 and PAX3 in the melanogenic pathway, it is possible that PAX3 regulates *Trp-1* expression, although an *in vivo*-based approach, such as ChIP, would be required to determine if this occurs through direct binding of PAX3 to elements in the *Trp-1* promoter.

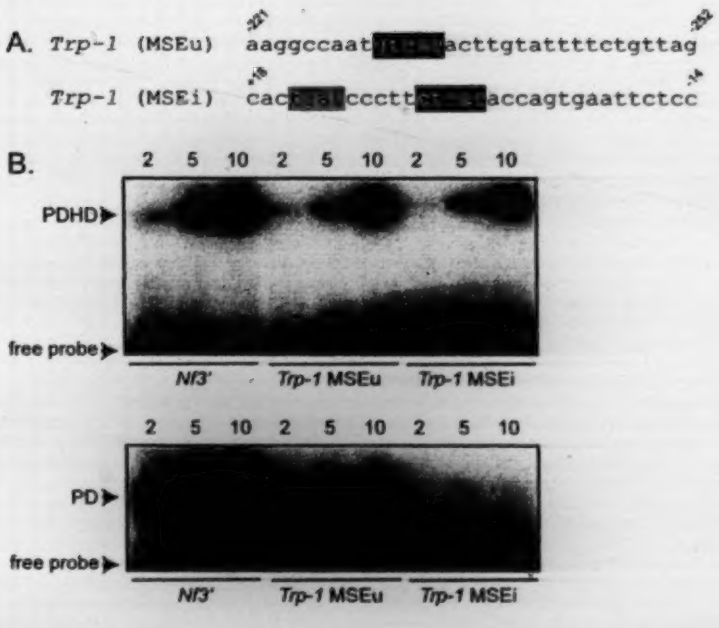

A. Trp-1 (MSEu) aaggccaat[illegible]acttgtattttctgttag
Trp-1 (MSEi) cac[illegible]ccctt[illegible]accagtgaattctcc
B.
2 5 10 2 5 10 2 5 10
PDHD
free probe
Nf3'
Trp-1 MSEu
Trp-1 MSEi
2 5 10 2 5 10 2 5 10
PD
free probe
Nf3'
Trp-1 MSEu
Trp-1 MSEi

Figure 3-5. Analysis of PAX3 binding to the *Trp-1* promoter. (A) The *Trp-1* promoter element contains two putative functional PAX3-responsive sites, MSEu and MSEi (see Section 3.6). Numbering indicates positioning with respect to the transcription start site. (B) PAX3 PDHD (*top*) and PD (*bottom*) binding to the MSEu and MSEi sites was analyzed using mobility shift assays. Positions of protein-DNA complexes and free probe are indicated.

3.7 Sequence determinants for *MITF* promoter binding by PAX3

The observation that tethering the HD to the PD improves PAX3 binding to the *MITF* and *MYF5* promoter sequences (Fig. 3-4) suggests that a cooperative binding event involving both domains mediates high affinity binding to composite elements. Based on the ability of the Prd PD and HD to bind to the juxtaposed recognition motifs of the *PH0* sequence (Jun and Desplan 1996), we predict that PAX3 interacts with the *MYF5* promoter sequence in a similar fashion. PAX3 recognition of the *Mitf* and *MITF* promoters, however, is a more complex scenario – both sequences feature multiple putative HD-recognition motifs (Fig. 3-1B) and our results so far do not clarify which site(s) is used. We therefore performed a series of DNA binding assays to clarify how the PAX3 HD interacts with the *MITF* promoter sequence. We first sought to determine which region of the *MITF* promoter element was responsible for mediating interaction with the HD and identified four potential HD recognition elements (hd1/2, hd3, and hd4) (Fig. 3-6A). The results of Watanabe et al. (1998) suggested that the overlapping HD motifs (hd1/2; 5'-ATTAAT) were important for *MITF* regulation by PAX3, but did not address whether they were necessary or sufficient for this function. Although these motifs resemble canonical HD-binding sites, they also overlap with the region predicted to bind the carboxy-terminal PD subdomain (see Fig. 3-1C), suggesting that PD binding would preclude interaction between the HD and these sites. The site adjacent to the PD motif (hd3) is a non-canonical site and is not conserved between mouse and human (see Fig. 3-2A), while the fourth site (hd4), also a non-canonical HD site, occupies the same position as the HD-binding motif of *e5* (although this motif is on the other strand in *e5*).

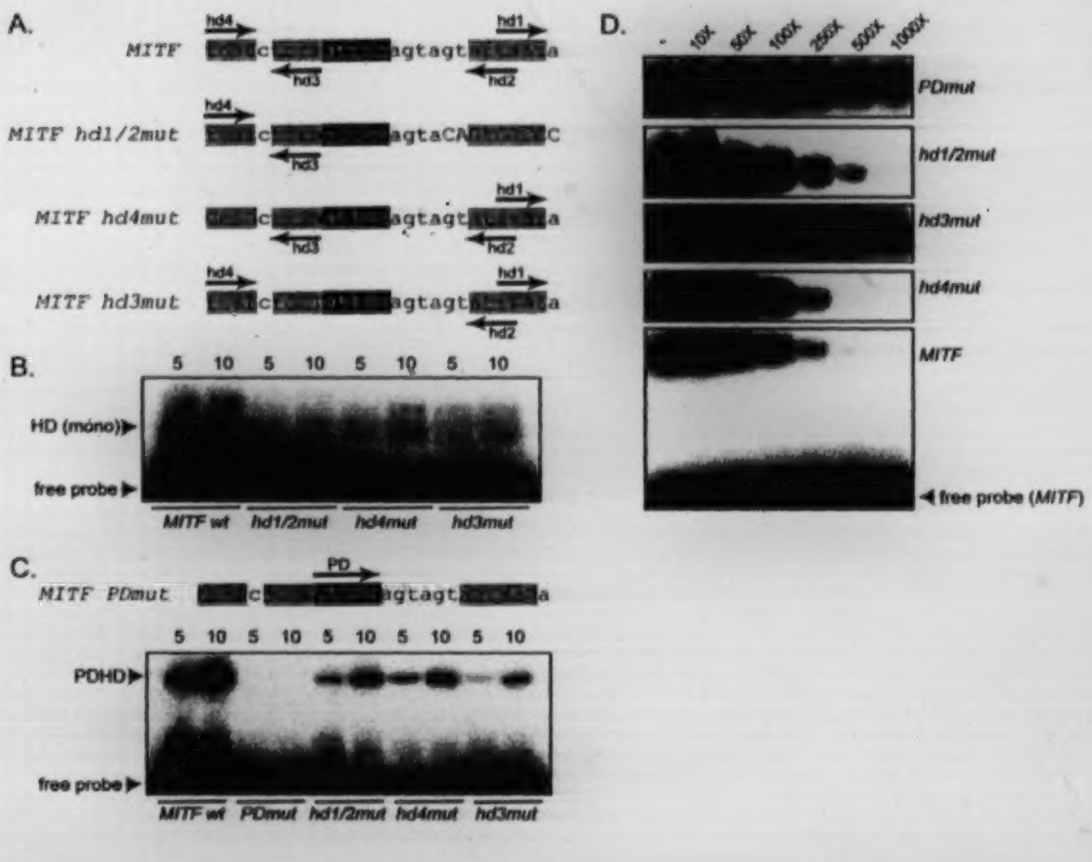

A.
hd4
hd1
MITF
c
agtagt
a
hd3
hd2
MITF hd1/2mut
agtaCA
C
MITF hd4mut
MITF hd3mut
B.
5 10 5 10 5 10 5 10
HD (mono)
free probe
MITF wt
hd1/2mut
hd4mut
hd3mut
C.
PD
MITF PDmut
5 10 5 10 5 10 5 10 5 10
PDHD
free probe
MITF wt
PDmut
hd1/2mut
hd4mut
hd3mut
D.
-
10X
50X
100X
250X
500X
1000X
PDmut
hd1/2mut
hd3mut
hd4mut
MITF
free probe (MITF)

Figure 3-6. Both the PAX3 paired domain and homeodomain make important contributions to *MITF* promoter binding. (A) The wild type *MITF* promoter sequence (*top*), showing locations of each putative HD-recognition motif, and sequences of *MITF* oligonucleotides in which the putative HD-recognition motifs have been mutated. Base changes are indicated by capital letters. The 5'-GTTCC PD motif is highlighted in black; putative HD motifs are highlighted in grey. (B) PAX3 HD binding to wild type and mutant *MITF* sequences. Position of HD-DNA complexes and free probe are indicated; amounts (μL) of protein added are shown above each lane. (C) PAX3 PDHD binding to wild type and mutant *MITF* sequences. Position of PDHD-DNA complexes and free probe are indicated; amounts (μL) of protein added are shown above each lane. HD mutant oligonucleotides are the same as in (A) and the *MITF PDmut* sequence is also shown. (D) Relative affinity of the PAX3 PDHD protein for the wild type and mutant *MITF* sequences was established using competition mobility shift assays. Free probe (wild type *MITF*) is shown for the bottom panel only; PDHD protein was used at a fixed concentration in all panels shown.

We performed mobility shift assays using the HD and oligonucleotides in which the hd1/2 sites were simultaneously mutated or in which the two upstream non-canonical sites were mutated individually (see above). Not surprisingly, our results demonstrated that mutating the overlapping canonical HD sites was most detrimental to PAX3 HD binding (Fig. 3-6B). We next examined the role of each HD binding site using the PDHD protein and the same HD-mutant oligonucleotides, as well as one containing a mutation of the GTTCC PD motif (Fig. 3-6C, *top*). PDHD binding was not affected by mutation of the downstream hd1/2 sites or the hd4 site (Fig. 3-6C); however, binding to the PD-mutated oligonucleotide was completely abrogated and interaction with the hd3 mutant oligonucleotide was noticeably reduced. Competition assays using the mutant oligonucleotides confirm the requirement for both the GTTCC and hd3 motifs (Fig. 3-6D), demonstrating that these are the primary sites of interaction for the PD and HD, respectively. Two important conclusions can be drawn from these experiments: first, PD binding is absolutely required for PAX3 interaction with the *MITF* sequence, and second, the hd3 motif that lies adjacent to the PD site appears to be the primary site of HD interaction, since affinity is decreased when this motif is mutated. It is important to note that the PDHD bound to all the HD mutant oligonucleotides, indicating some flexibility in HD binding.

3.8 Sequence requirements within the *MITF* hd3 motif

Our results suggest that, despite being a non-canonical Prd-type HD recognition site, the hd3 motif is required for high affinity binding of PAX3 to the *MITF* promoter sequence. To clarify how the HD interacts with this motif, we mutated each base of the hd3 site to guanine individually and evaluated PDHD binding to these oligonucleotides (Fig. 3-7). The hd1/2 motifs were simultaneously mutated to eliminate any confusion caused by the binding of multiple HDs. The *MITF gaaa* oligonucleotide, which is analogous to the *Mitf* promoter sequence (*cf.* Fig. 3-1B), competed with an efficiency equivalent to that of the wild type *MITF* oligonucleotide (Fig. 3-7A, *second panel*). In contrast, mutating the middle two adenines of the 5'-TAAA motif severely reduced competition for PDHD binding (Fig. 3-7A, *third & fourth panels*). Finally, mutating the hd3 motif to 5'-TAAG did not appreciably affect PDHD binding, as competition efficiency was only slightly below that of the *MITF* oligonucleotide (Fig. 3-7A, *bottom panel*).

PDHD binding to labeled mutant oligonucleotides displayed a similar trend and made the effects of the mutations more obvious (Fig. 3-7B). The PDHD showed maximal affinity for the *MITF hd3opt* sequence, which contains an optimized hd3 motif, and affinity for the *MITF gaaa* sequence was also high, reflecting the result of the competition assays above. Also correlating with the competition assays, PDHD affinity for the *MITF tgaa* and *MITF taga* oligonucleotides was reduced. We observed minimal binding to the *MITF taga* sequence, suggesting the middle adenine of 5'-TAAA plays an important role in HD contact. Binding to the *MITF taag* sequence was slightly reduced compared to *MITF hd3opt*, which is also in accordance with the competition assay results. Together, these data indicate that the AA-dinucleotide that forms the central core of the 5'-TAAA motif is critical for HD-binding and PD-HD cooperativity, with the most important base pair being the second adenine-thymine (TAAA). Our results also demonstrate that the PAX3 HD exhibits a remarkable degree of flexibility in binding to

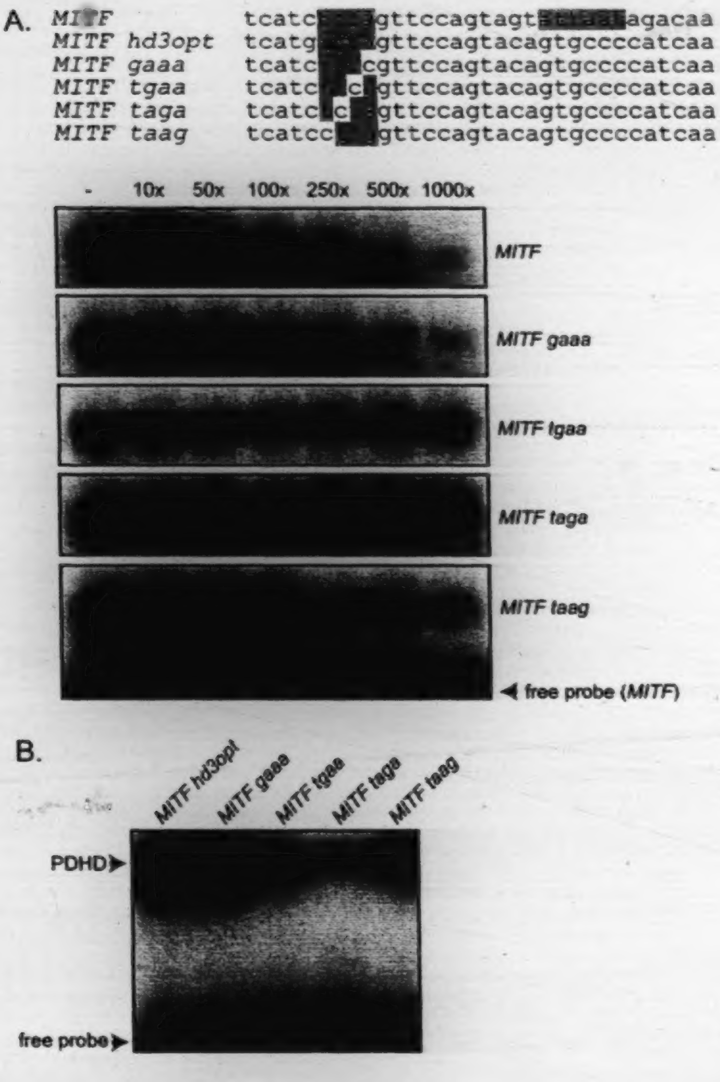

A.
MITF tcatc[illegible]gttccagtagt[illegible]agacaa
MITF hd3opt tcatg[illegible]gttccagtacagtgccccatcaa
MITF gaaa tcatc[illegible]cgttccagtacagtgccccatcaa
MITF tgaa tcatc[illegible]c[illegible]gttccagtacagtgccccatcaa
MITF taga tcatc[illegible]c[illegible]gttccagtacagtgccccatcaa
MITF taag tcatcc[illegible]gttccagtacagtgccccatcaa
-
10x
50x
100x
250x
500x
1000x
MITF
MITF gaaa
MITF tgaa
MITF taga
MITF taag
free probe (MITF)
B.
MITF hd3opt
MITF gaaa
MITF tgaa
MITF taga
MITF taag
PDHD
free probe

Figure 3-7. Analysis of the *MITF* hd3 motif. (A) *Top*, alignment of the wild type *MITF* oligonucleotide and derivatives containing optimized (*MITF hd3opt hd1/2mut*) or mutated hd3 motifs (all derivative oligonucleotides contain mutated hd1/2 motifs). Putative HD-binding sites are highlighted in grey. *Bottom*, competition mobility shift assays were used to evaluate the effects of mutating individual bases in the hd3 motif. Free probe (wild type *MITF*) is shown for the bottom panel only; PDHD protein was used at a fixed concentration in all panels shown. (B) PDHD affinity for the hd3-optimized and mutant sequences was assessed in standard mobility shift assays.

TAAT-like sequences, and show that a canonical HD motif is not necessary for efficient DNA binding.

3.9 Structural requirements for PAX3-*MITF* interaction

So far, our results suggest that PAX3 interaction with the *MITF* promoter sequence is a PD-driven event that requires participation of the HD for high affinity binding. To further examine the structural determinants of PAX3 binding to the *MITF* sequence and to assess PD-HD cooperativity, we performed recruitment assays with isolated PD and HD to see if a ternary complex including the domains and the *MITF* oligonucleotide would be formed. Indeed, we observed a ternary complex on the *MITF* sequence (Fig. 3-8A, *top panel*), demonstrating that the two domains need not be linked for efficient binding. We also examined PD and HD binding to the mouse *Mitf* promoter sequence, which features an alteration in the hd3 motif (5'-TAAA→GAAA). Despite this change, we observed efficient PD and HD binding and formation of a ternary complex (Fig. 3-8A, *bottom panel*); however, the dimeric HD complex formed on the human *MITF* oligonucleotide was no longer observed in the presence of the *Mitf* oligonucleotide, suggesting that a 5'-GAAA motif is insufficient to mediate HD interaction in the absence of the PD. More importantly, these data show that the base immediately adjacent to the PD motif is not essential for efficient cooperative binding, consistent with its lack of conservation among higher eukaryotes (Fig. 3-2A). To validate the importance of the hd3 site for efficient PAX3 binding, we performed a recruitment assay in the presence of an *MITF* oligonucleotide containing a destroyed hd3 site to see if a ternary complex could still be formed. We observed PD binding and

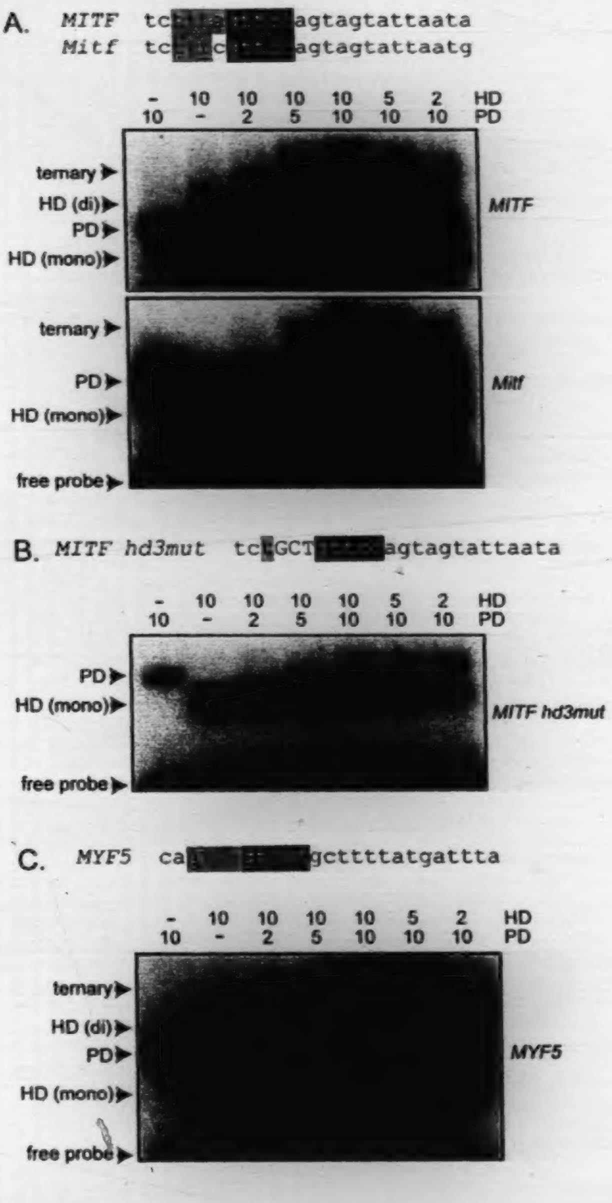

A.
MITF tc agtagtattaata
Mitf tc c agtagtattaatg
- 10 10 10 10 5 2 HD
10 - 2 5 10 10 10 PD
ternary
HD (di)
PD
HD (mono)
MITF
ternary
PD
HD (mono)
free probe
Mitf
B. MITF hd3mut tc GCT agtagtattaata
- 10 10 10 10 5 2 HD
10 - 2 5 10 10 10 PD
PD
HD (mono)
free probe
MITF hd3mut
C. MYF5 ca gcttttatgattta
- 10 10 10 10 5 2 HD
10 - 2 5 10 10 10 PD
ternary
HD (di)
PD
HD (mono)
free probe
MYF5

Figure 3-8. Composite promoter elements support cooperative binding by the PAX3 paired domain and homeodomain. (A) Cooperative binding of the PAX3 PD and HD to the *MITF* (*top panel*) and *Mitf* (*bottom panel*) oligonucleotides was investigated using recruitment assays. Amounts (µL) of each protein used are indicated above each lane. Positions of the PD-DNA, HD-DNA (*mono* = monomer; *di* = dimer), and ternary complexes are indicated. Position of the free probe is shown for the bottom panel only. (B) No ternary complex is formed on an oligonucleotide containing a mutation of the hd3 motif (*MITF hd3mut*). Mutated bases are capitalized. (C) The PAX3 binding site of the *MYF5* promoter supports cooperative binding by the PAX3 PD and HD.

monomeric HD binding to the oligonucleotide, but no ternary complex was formed (Fig. 3-8B). Interestingly, no dimeric HD complex was observed, suggesting that the hd3 motif, although not an optimal HD site, acts as a secondary site for HD binding.

Recently, it was demonstrated that PAX3 could regulate expression of *Myf5* through a binding site in the gene's promoter (Bajard et al. 2006). Similar to the *PH0* sequence (Jun and Desplan 1996), both the *Myf5* and *MYF5* promoters contain a putative HD-binding consensus (5-ATTA) adjacent to the PD motif (Fig. 3-1B), suggesting that the sequences could mediate a binding event involving the PD and HD, similar to *MITF*. As shown above for *MITF* and *Mitf*, the PAX3 PD and HD efficiently formed a ternary complex on the *MYF5* (Fig. 3-8C) and *Myf5* (not shown) sequences. In their study, Bajard et al. (2006) mutated the *MYF5* binding site and demonstrated that PAX3 no longer bound when the PD site was destroyed. However, the mutation also altered the adjacent 5'-ATTA motif, and in doing so, failed to provide information on the importance or relevance of HD binding to the *MYF5* sequence. We individually mutated the PD and HD sites of the *MYF5* sequence and performed recruitment binding assays with the PD and HD, but observed no ternary complex formation in the presence of either oligonucleotide (not shown), suggesting that, like *MITF*, binding by both the PD and HD is important for PAX3 interaction with the *MYF5* promoter. In support, a recent study investigating effects of *Myf5* promoter mutations demonstrated that both the 5'-ATTA and 5'-GTCAT motifs were critical for PAX3 binding (Buchberger et al. 2007).

3.10 Participation of paired domain and homeodomain homologs in cooperative binding

All PAX family members contain a PD but feature different combinations of the other conserved regions, including the octapeptide and HD (Fig. 1-2). So far, we have examined the DNA binding behavior of the PAX3 PD alone, in the context of the full length protein, and its cooperative binding properties with the HD. Since the PAX3 PD was able to recruit the PAX3 HD to the composite *MITF* and *MYF5* sequences, we wished to investigate whether this property was shared among PAX family PDs. The extensive sequence homology among all PDs suggests that their overall behavior in regards to DNA interaction should be similar, although not all PDs bind preferentially to the same recognition sequence. For example, the PAX1 PD binds preferentially to an *e5*-like site (Chalepakis et al. 1991), while the PAX2 and PAX3 PDs bind optimally to sequences featuring a 5'-GTCAC motif, and the PAX6 PD prefers sequences with a 5'-TTCAC motif (Epstein et al. 1994; Epstein et al. 1995). These observations imply that, despite their extensive sequence identity, PDs of different classes contain specific determinants that dictate their DNA binding specificity and sequence preference. To address the ability of various PAX PDs to support ternary complex formation, we performed recruitment assays with representative PDs from each PAX family class (Fig. 3-9A) and the *MITF* oligonucleotide. This allowed us to examine PD binding on its own and to assess whether the PDs can participate in a cooperative binding event with the PAX3 HD. These experiments also provide a basis for identifying determinants for cooperativity contained in the PD. The PAX1, PAX2, and PAX6 PDs each bound to the *MITF* sequence and all formed a ternary complex with PAX3 HD (Fig. 3-9B),

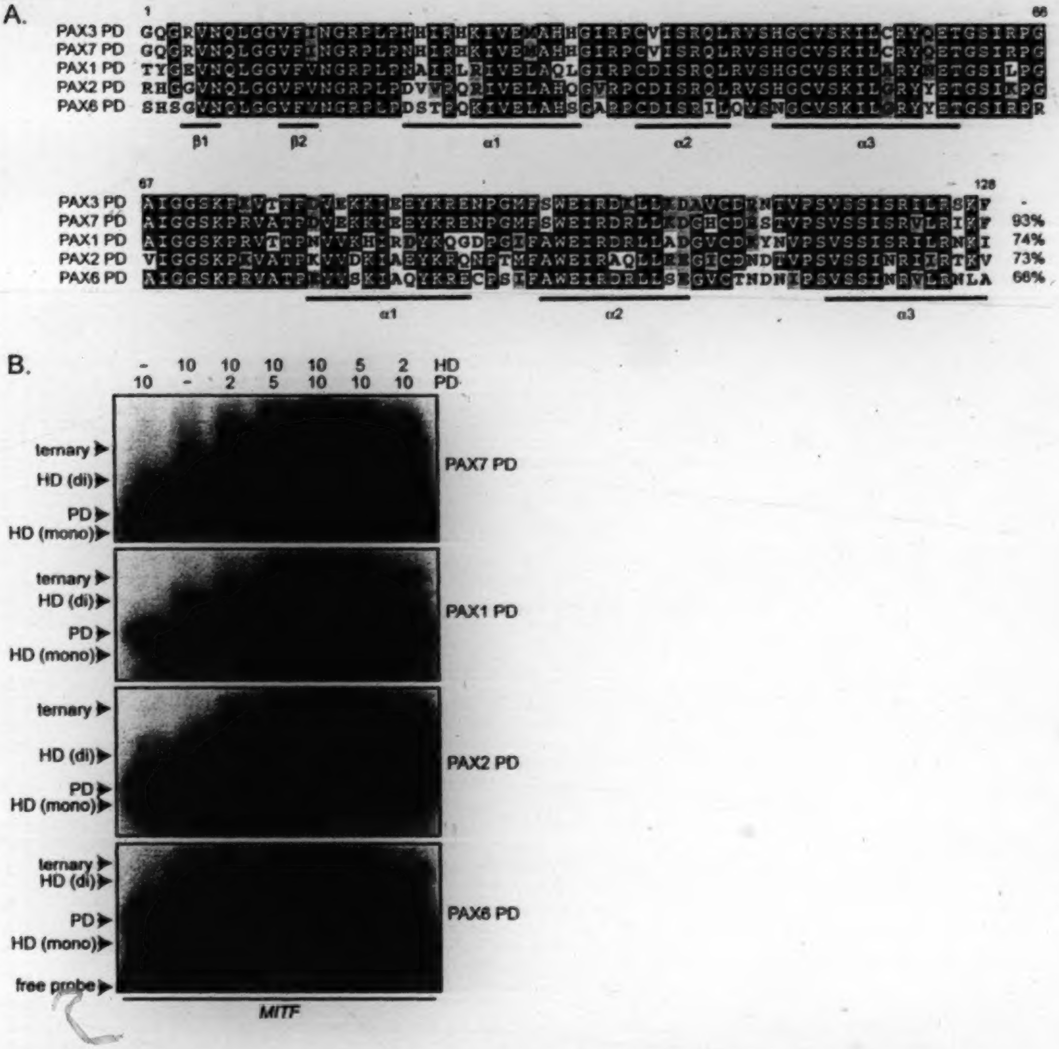

A.
1
66
PAX3 PD GQGRVNQLGGVFINGRPLPNHIRHKIVEMAHHGIRPCVISRQLRVSHGCVSKILCRYQETGSIRPG
PAX7 PD GQGRVNQLGGVFINGRPLPNHIRHKIVEMAHHGIRPCVISRQLRVSHGCVSKILCRYQETGSIRPG
PAX1 PD TYGEVNQLGGVFVNGRPLPNAIRLRIVELAQLGIRPCDISRQLRVSHGCVSKILARYNETGSILPG
PAX2 PD RHGGVNQLGGVFVNGRPLPDVVRQRIVELAHQGVRPCDISRQLRVSHGCVSKILGRYYETGSIKPG
PAX6 PD SHSGVNQLGGVFVNGRPLPDSTRQKIVELAHSGARPCDISRILQVSNGCVSKILGRYYETGSIRPR
β1
β2
α1
α2
α3
67
128
PAX3 PD AIGGSKPKVTTPDVEKKIEEYKRENPGMFSWEIRDKLLKDAVCDRNTVPSVSSISRILRSKF -
PAX7 PD AIGGSKPRVATPDVEKKIEEYKRENPGMFSWEIRDRLLKDGHCDRSTVPSVSSISRVLRIKF 93%
PAX1 PD AIGGSKPRVTTPNVVKHIRDYKQGDPGIFAWEIRDRLLADGVCDKYNVPSVSSISRILRNKI 74%
PAX2 PD VIGGSKPKVATPKVVDKIAEYKRQNPTMFAWEIRAQLLREGICDNDTVPSVSSINRIIRTKV 73%
PAX6 PD AIGGSKPRVATPEVVSKIAQYKRECPSIFAWEIRDRLLSEGVCTNDNIPSVSSINRVLRNLA 66%
α1
α2
α3
B.
- 10 10 10 10 5 2 HD
10 - 2 5 10 10 10 PD
ternary
HD (di)
PD
HD (mono)
PAX7 PD
PAX1 PD
PAX2 PD
PAX6 PD
free probe
MITF

Figure 3-9. Paired domain paralogs can recruit the PAX3 homeodomain to the *MITF* promoter sequence. (A) Alignment of PD amino acid sequences from the PAX3, PAX7, PAX1, PAX2, and PAX6 proteins. Identical residues are shaded in black, conservative substitutions are shaded in gray. Percent identity to the PAX3 PD is indicated to the right of each sequence. Sequences were aligned using ClustalW and shaded with BoxShade 3.21. (B) Cooperative binding of the PAX7 (*top panel*), PAX1 (*second panel*), PAX2 (*third panel*), and PAX6 (*bottom panel*) PDs and the PAX3 HD to the wild type *MITF* sequence was investigated using recruitment assays. Amounts (μL) of each protein used are indicated above each lane. Positions of the PD-DNA, HD-DNA (*mono* = monomer; *di* = dimer), and ternary complexes are indicated. Position of the free probe is shown for the bottom panel only.

demonstrating that members of each PAX subclass are able to interact with GTTCC-type PD motifs. Surprisingly, these results also show that PDs from each class, including those that do not feature a HD, can cooperatively bind to composite PD-HD recognition motifs.

We also tested the ability of the PAX3 PD to recruit a diverse selection of HDs to the *MITF* promoter element. All HDs share the HTH structure and almost all bind to a common recognition sequence defined by a core 5'-TAAT motif (Gehring et al. 1994a). We chose representative HDs from PAX family proteins that feature a HD (Fig. 3-10A): the PAX7 HD, which shares the most extensive sequence identity with PAX3, the PAX6 HD, which shares less identity with PAX3, and the PAX2 rudimentary HD, which lacks the HD HTH motif. We also examined the Prd-type PRRX1 (Grueneberg et al. 1992) and PITX2 (Semina et al. 1996) HDs, as well as the HD from MSX1, a HOX class homeoprotein (Hill et al. 1989). Interestingly, the only HD to form a ternary complex on the *MITF* oligonucleotide was the PAX7 HD (Fig. 3-10B, *top panel*). As expected, the rudimentary PAX2 HD failed to bind to *MITF* (Fig. 3-10B, *second panel*). The PAX6, PRRX1, and PITX2 HDs all bound to the *MITF* sequence, but none permitted formation of a ternary complex (Fig. 3-10B, *panels 3-5*). The MSX1 HD bound with low affinity as a monomer and dimer, but did not form a ternary complex (Fig. 3-10B, *bottom panel*). Importantly, these results suggest that the PAX3 PD can cooperatively bind to composite sequences only with HDs that share a high degree of sequence identity with the PAX3 HD, such as the PAX7 HD. Together, these results suggest that important differences among Prd-class HDs modulate their behavior and that, among these HDs, the

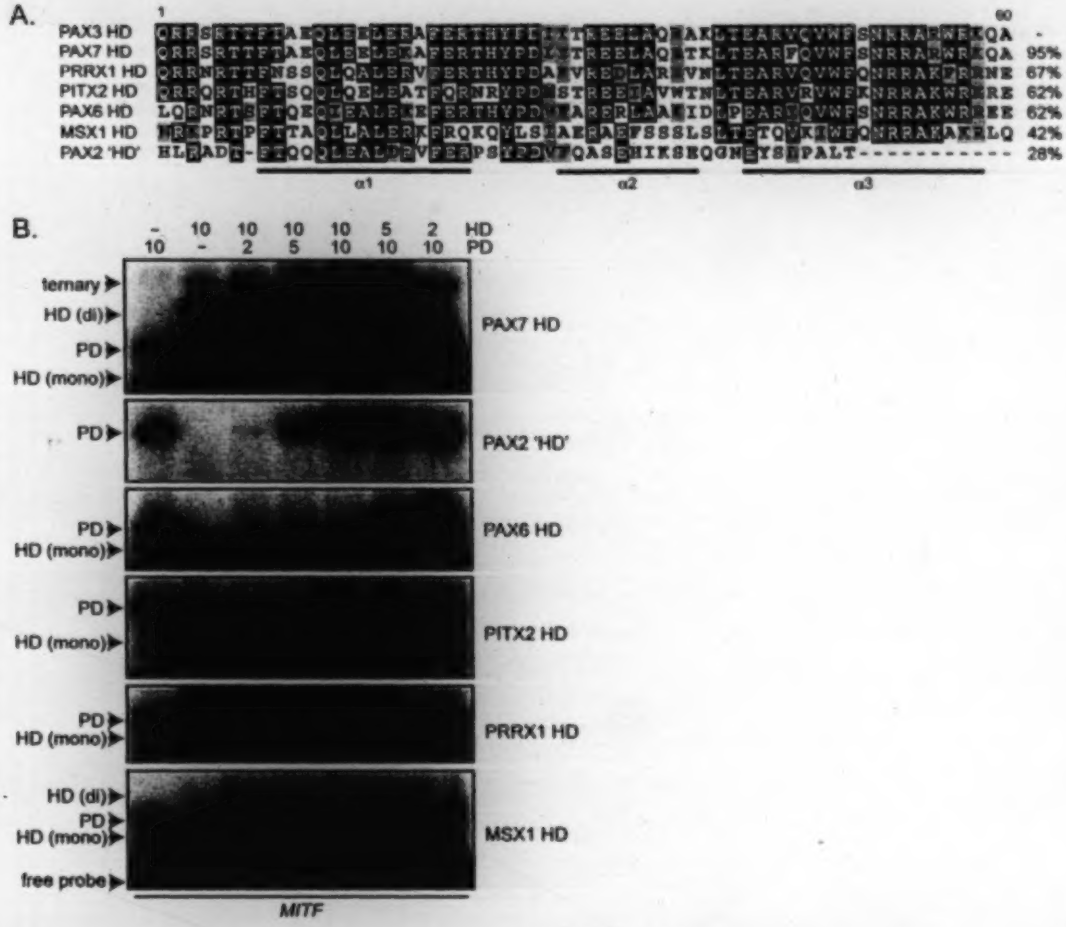

A.
1
60
PAX3 HD
PAX7 HD
PRRX1 HD
PITX2 HD
PAX6 HD
MSX1 HD
PAX2 'HD'
-
95%
67%
62%
62%
42%
28%
α1
α2
α3
B.
- 10 10 10 10 5 2 HD
10 - 2 5 10 10 10 PD
ternary
HD (di)
PD
HD (mono)
PAX7 HD
PAX2 'HD'
PAX6 HD
PITX2 HD
PRRX1 HD
MSX1 HD
free probe
MITF

Figure 3-10. PAX3/7 homeodomain homologs cannot cooperatively bind to a composite sequence with the PAX3 paired domain. (A) Alignment of HD amino acid sequences from the PAX3, PAX7, PRRX1, PITX2, PAX6, and MSX1 proteins, as well as the rudimentary HD from PAX2. Identical residues are shaded in black, conservative substitutions are shaded in gray. Percent identity to the PAX3 HD is indicated to the right of each sequence. Sequences were aligned using ClustalW and shaded with BoxShade 3.21. (B) Cooperative binding of the HD homologs and the PAX3 PD to the wild type *MITF* sequence was investigated using recruitment assays. Amounts (µL) of each protein used are indicated above each lane. Positions of the PD-DNA, HD-DNA (*mono* = monomer; *di* = dimer), and ternary complexes are indicated. Position of the free probe is shown for the bottom panel only.

determinants responsible for dictating functional interaction are contained in the HD, not in the PAX3 PD.

3.11 Effects of disease mutations on PAX3 target recognition

PAX3 mutations, most of which occur in the PD and HD, underlie WS1 and WS3 in humans and the murine *Sp* phenotype (see Section 1.1.5). DNA binding assays using optimized binding sites for the PD and HD (*Nf3'* and *P2*, respectively) demonstrate that mutations in one domain typically affect binding by the other, suggesting that the PD and HD functionally interact (Underhill et al. 1995; Fortin et al. 1997; Underhill and Gros 1997). However, the influence of these mutations on *in vivo*-derived target site recognition is unknown, as is their effect on binding to composite sequences. To resolve this, we analyzed a subset of WS mutations (F45L, N47H, G81A, S84F, and Y90H in the PD, and V265F and R271G in the HD), along with the Sp^d mutation (G42R), to determine their impact on DNA binding in the context of the recombinant PDHD protein (see Fig. 3-11A for the location of these mutations in the PDHD protein).

We first examined the affinity of mutant PDs for the *MITF* and *Trp-1* (MSEu) sites (Fig. 3-11B). As before, the wild type PD bound to both sites. The N47H PD displayed enhanced binding to both sites, but since the full length N47H mutant failed to interact with the PD-binding *Nf3'* and *H2A2.1* sequences (Fortin et al. 1997), other regions of PAX3 may influence PD-DNA interaction in this particular mutant. In the case of the G81A PD, we observed limited binding to *Trp-1* relative to wild type, despite no detectable binding to *MITF*, suggesting that this mutation affects discrimination of disparate recognition motifs by the PAX3 PD. The specificity of these effects is

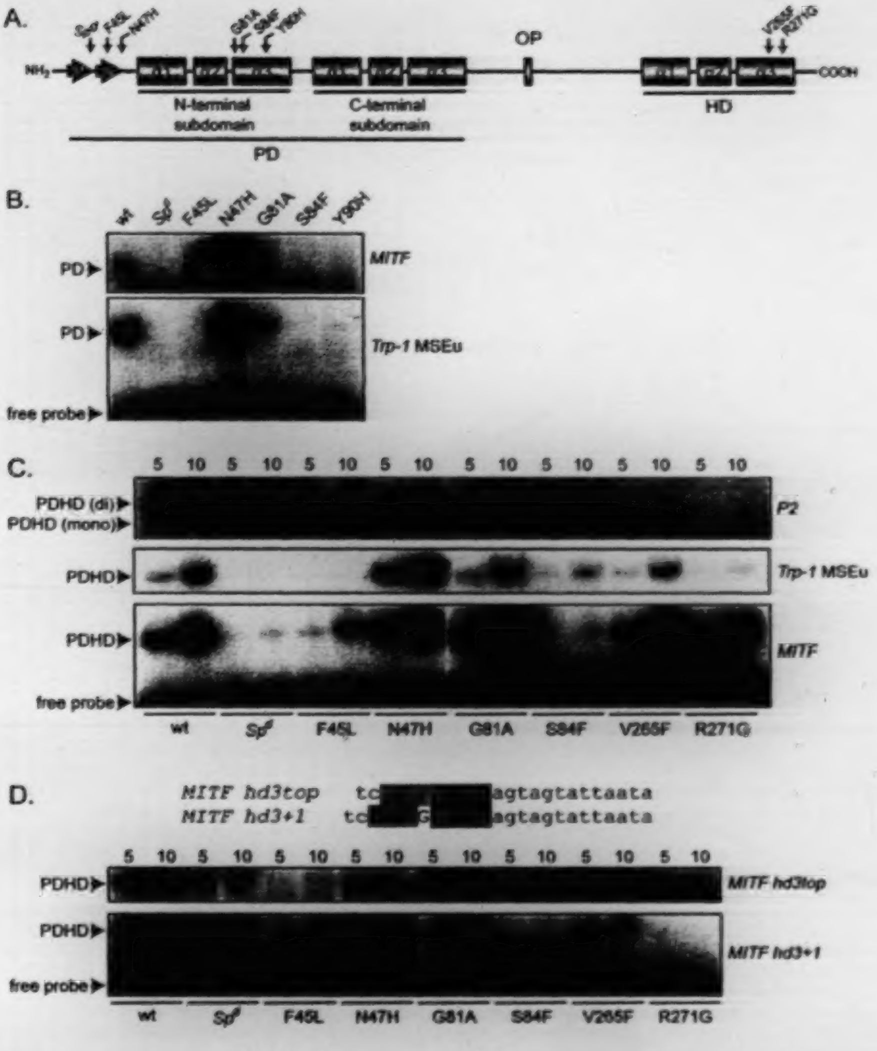

A.
Sp^d
F45L
N47H
G81A
S84F
Y90H
OP
V265F
R271G
NH2
COOH
N-terminal subdomain
C-terminal subdomain
HD
PD
B.
wt
Sp^d
F45L
N47H
G81A
S84F
Y90H
PD
MITF
PD
Trp-1 MSEu
free probe
C.
5 10 5 10 5 10 5 10 5 10 5 10 5 10 5 10
PDHD (di)
PDHD (mono)
P2
PDHD
Trp-1 MSEu
PDHD
MITF
free probe
wt
Sp^d
F45L
N47H
G81A
S84F
V265F
R271G
D.
MITF hd3top tc agtagtattaata
MITF hd3+1 tc G agtagtattaata
5 10 5 10 5 10 5 10 5 10 5 10 5 10 5 10
PDHD
MITF hd3top
PDHD
MITF hd3+1
free probe
wt
Sp^d
F45L
N47H
G81A
S84F
V265F
R271G

Figure 3-11. Effects of disease mutations on PAX3 DNA binding behavior. (A) Schematic representation of the PDHD protein (not to scale), showing the location of mutations analyzed in DNA binding assays. OP, octapeptide. (B) Wild type and mutant PD binding to the *MITF* (*top panel*) and *Trp-1* MSEu (*bottom panel*) oligonucleotides. Positions of PD-DNA complexes and free probe (bottom panel only) are indicated; all protein concentrations were equalized to that of wild type PD. (C) Wild type and mutant PDHD binding to the *P2* (*top panel*), *Trp-1* MSEu (*middle panel*), and *MITF* (*bottom panel*) oligonucleotides. Positions of PDHD-DNA complexes and free probe (bottom panel only) are indicated; amounts (μL) of protein added are shown above each lane. All protein concentrations were equalized to that of wild type PDHD. (D) *Top panel*, PDHD binding to the *MITF hd3top* oligonucleotide. *Bottom panel*, PDHD binding to the *MITF hd3+1* oligonucleotide, which contains a one bp insertion between the hd3 motif and GTTCC PD motif. All protein concentrations were equalized to that of wild type PDHD. Oligonucleotide sequences are shown above.

established by the fact the remaining mutants (*Sp*^{d}, F45L, S84F, and Y90H) failed to bind either site.

Next, we performed mobility shift assays with wild type and mutant PDHD proteins to investigate effects of the PD and HD mutations on PAX3 binding in the context of both domains. We used three different oligonucleotides to assess the effects of mutations on PAX3-DNA interaction: *P2*, which confers dimeric binding by the HD, *Trp-1*, which represents a PD-only binding site, and *MITF*, a composite PD-HD sequence (Fig. 3-1). The PDHD protein efficiently formed a dimeric complex on the *P2* sequence (Fig. 3-11C, *top panel*), in accordance with prior investigations of the full length PAX3 protein (Underhill et al. 1995). The N47H and G81A mutants bound efficiently to *P2* as dimers and we observed low levels of monomeric binding by the *Sp*^{d} and F45L PDHDs. Both the S84F and R271G mutants failed to bind to the *P2* sequence, while the V265F mutant displayed reduced levels of dimeric binding. The latter result suggests, as others have shown (Underhill et al. 1995; Underhill and Gros 1997), that the presence of the PD can influence HD behavior, since the V265F HD fails to bind to the *P1/2* or *P2* sequences in the absence of the PD (not shown). The N47H mutant retained the highest level of affinity for the *Trp-1* sequence, while the G81A and V265F mutants displayed reduced but detectable binding (Fig. 3-11C, *middle panel*). We observed no interaction between the *Sp*^{d} and F45L mutants and the *Trp-1* oligonucleotide, while the S84F mutant displayed considerably reduced binding levels. The R271G mutant did not bind to the *Trp-1* sequence, which, together with the results of Fortin et al. (1997) serves to show that a mutation in the HD can have a deleterious effect on PAX3 interaction with PD-specific sequences. In contrast to our observations using the *Trp-1* oligonucleotide, the

F45L and R271G mutants bound to the *MITF* sequence (Fig. 3-11C, *bottom panel*), even though the corresponding PD (Fig. 3-11B) and HD (see below) proteins displayed undetectable DNA binding activity. The R271G mutant bound to *MITF* at comparable levels to the wild type PAX3 PDHD, but the *Sp*^{d} and S84F mutants were still deficient in DNA binding. Finally, the Y90H PDHD bound to the *Trp-1* and *MITF* sequences with affinity comparable to the wild type PDHD (not shown). These data indicate that combinatorial use of the PD and HD can restore DNA binding activity to mutant proteins in a binding-site dependent manner.

PAX3 binding to the *MITF* site is context dependent, as demonstrated by the effects of reversing the polarity (*MITF hd3top*) or spacing (*MITF hd3+1*) of the hd3 motif with respect to the PD binding site (Fig. 3-11D). Altering the location of the hd3 motif did not hinder binding by the wild type PDHD protein, consistent with flexibility of PAX3 on the *e5* element (Chalepakis et al. 1994c); however, the F45L and R271G mutations displayed reduced or absent binding to each sequence. In addition to altered DNA binding activity, it is clear that there are differences in binding specificity for each of the mutants tested: N47H bound less efficiently than wild type PDHD to the *MITF hd3top* element, while G81A and V265F exhibited this trend on the *MITF hd3+1* site (*cf.* Figs. 3-11C & D). Importantly, these data indicate that *in vitro*-derived binding sites may not necessarily provide an accurate reflection of the DNA binding defects caused by PAX3 disease mutations. These analyses also establish that individual mutations show considerable disparity in their effects on specific target sites, supporting the use of different DNA binding mechanisms by PAX3.

3.12 Effects of PAX3 disease mutations on paired domain and homeodomain cooperativity

Our results demonstrate that disease mutations exert a range of effects on PAX3 interaction with HD-specific (*P2*), PD-only (*Trp-1* MSEu), and composite (*MITF*) binding sequences and suggest that inter-domain cooperativity is also affected by disease mutations. To assess effects on cooperativity, we performed recruitment assays in the presence of the *MITF* oligonucleotide, on which the wild type PD and HD efficiently form a ternary complex (see Fig. 3-8A). We first examined the ability of recombinant PD proteins containing the same mutations in Fig. 3-11 to form a ternary complex in the presence of the HD. As before, the wild type PD bound with moderate affinity to *MITF* and efficiently formed a ternary complex with the HD (Fig. 3-12A). The *Sp*^{d} and F45L PDs, which cannot bind to the *MITF* sequence on their own, did not form a ternary complex with the HD. In contrast, the N47H PD, which bound with high affinity to *MITF* in the absence of the HD, formed a ternary complex upon addition of the HD. Despite abrogating binding to the *MITF* sequence, the G81A mutation permitted formation of a faint ternary complex in the presence of the HD, suggesting that this mutant retains inter-domain cooperativity. Finally, we observed no ternary complex formation in the presence of the S84F or Y90H mutant PDs. Together, these results suggest that PD-HD cooperativity is an important determinant of PAX3 DNA binding ability, since the mutations that abrogated recruitment in this assay were the same ones that affected PDHD binding to the *MITF* and *Trp-1* sites above.

To examine mutant HD binding to the *MITF* sequence, and to investigate the effects of HD mutations on cooperative binding by the PD and HD, we performed

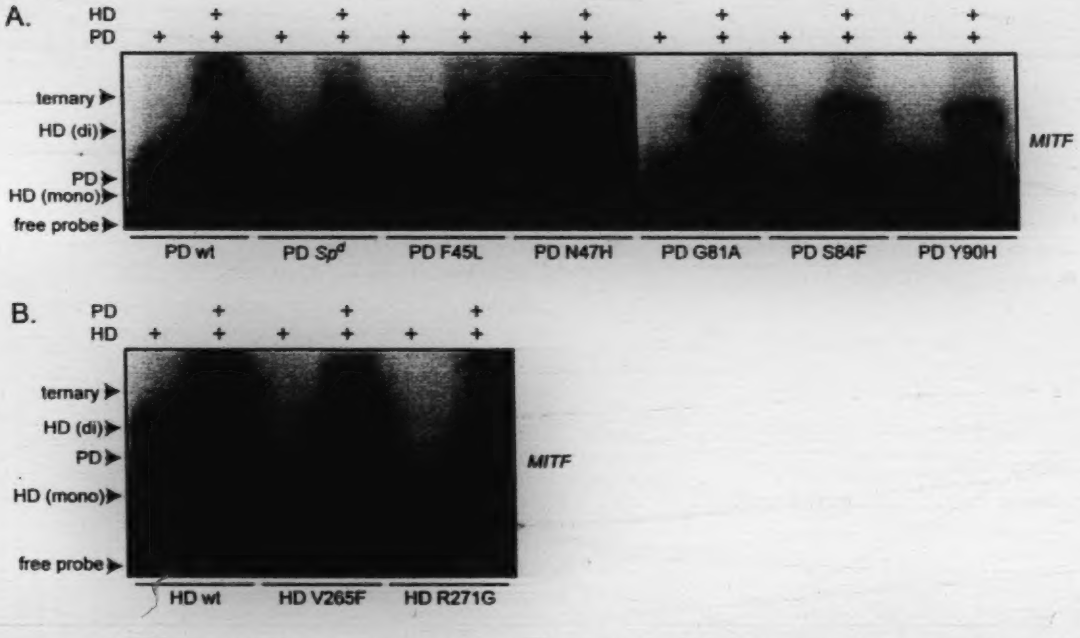

A.
HD + + + + + + +
PD + + + + + + + + + + + + + +
ternary
HD (di)
PD
HD (mono)
free probe
MITF
PD wt
PD Sp^{d}
PD F45L
PD N47H
PD G81A
PD S84F
PD Y90H
B.
PD + + +
HD + + + + + +
ternary
HD (di)
PD
HD (mono)
free probe
MITF
HD wt
HD V265F
HD R271G

Figure 3-12. PAX3 disease mutations interfere with inter-domain cooperativity. (A) Cooperative binding of mutant PD proteins and the wild type PAX3 HD was assessed using recruitment binding assays in the presence of the wild type *MITF* oligonucleotide. (B) Recruitment binding assays to examine effects of the V265F and R271G HD mutations on PD-HD cooperativity. Locations of protein-DNA complexes and free probe are indicated.

recruitment binding assays with the wild type PD and V265F and R271G HDs. Our results show that the V265F HD bound to the *MITF* sequence as a monomer, but DNA binding by the R271G HD was undetectable (Fig. 3-12B). Additionally, the V265F HD supported weak ternary complex formation in the presence of the PD, but the PD could not recruit the R271G HD to the *MITF* sequence, despite high affinity PD binding. These data show that, although the V265F mutation imparts a defect on HD-DNA interaction, this defect can be overcome by ternary complex formation in the presence of the PD. This may account for the fact that the V265F PDHD retained the ability to interact with the *Trp-1* and *MITF* sequences (see Fig. 3-11B). Conversely, the R271G mutation eliminates DNA binding by the isolated HD and appears to affect PD-DNA binding and functional cooperativity between the two domains, as demonstrated by the inability of the R271G PDHD to bind to the *Trp-1* sequence (Fig. 3-11B) and derivatives of the *MITF* site in which the hd3 motif has been repositioned (Fig. 3-11D).

3.13 Functional analysis of the *MITF* and *Trp-1* PAX3 recognition elements

The above results indicate that differential use of its DNA-binding domains allows PAX3 to interact with distinct recognition motifs in the *MITF* and *Trp-1* promoters. Specifically, binding to the *MITF* promoter requires involvement of both the PD and HD, while *Trp-1* binding is facilitated primarily by the PD. Our data also show that disease mutations affect PAX3 binding to the *Trp-1* and *MITF* elements to varying extents. To understand the functional implications of these observations, we evaluated the ability of wild type and mutant PAX3 to regulate transactivation of a luciferase reporter gene in the context of each promoter using a heterologous cell system. Two

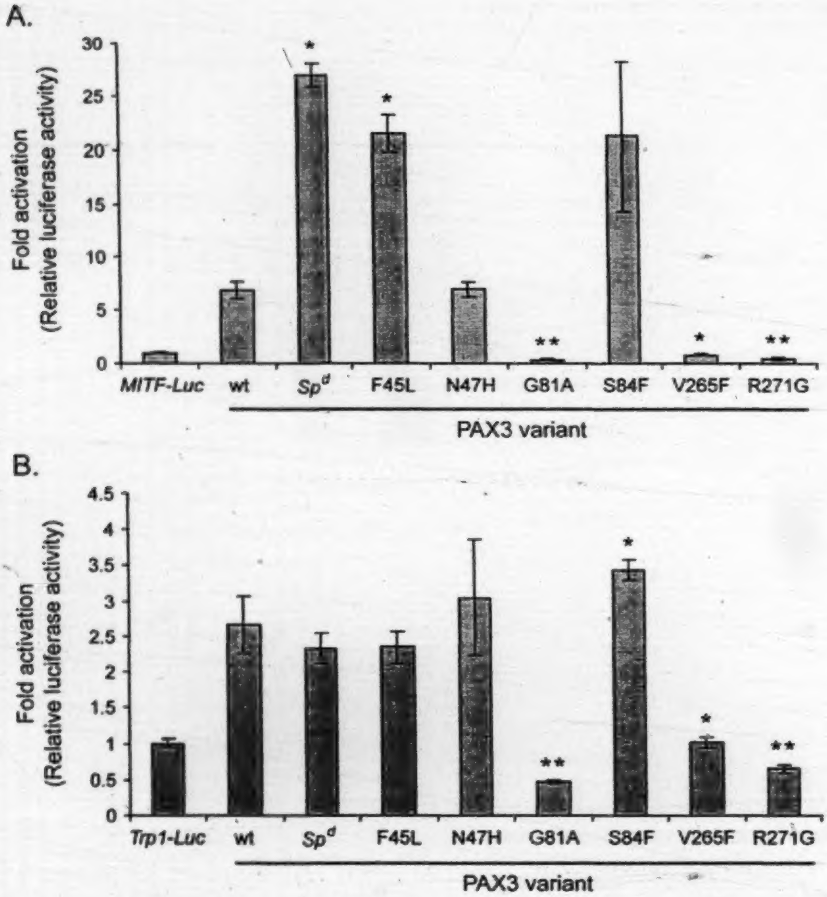


Figure 3-13. Functional analysis of the *MITF* and *Trp-1* PAX3-binding sites. Regulation of the *MITF-Luc* (A) or *Trp1-Luc* (B) reporter genes by wild type or mutant PAX3 was analyzed in a dual luciferase assay (see Section 2.8). HeLa cells were transfected with 0.5 µg of reporter plasmid DNA, 0.5 µg of PAX3 expression plasmid DNA, and 5 ng of the *Renilla* luciferase pRL-TK control plasmid. Whole cell lysates were assayed for luciferase activity 24 h post transfection. As a standard reference, activity of each reporter plasmid in the absence of a PAX3 expression plasmid was normalized to one relative luciferase unit; remaining values are expressed as fold-activation relative to the normalized reporter plasmid. The pGL3-basic vector was used as a negative control (not shown). Values represent the mean and standard deviation from one experiment, done in triplicate. One asterisk indicates significant levels of activation or repression compared to wild type PAX3; two asterisks indicate significant levels of repression compared to reporter plasmid alone and wild type PAX3.

reporter constructs were used, one containing the upstream regulatory sequence of the *MITF* gene (*MITF-Luc*), and the other containing the *Trp-1* promoter (*Trp1-Luc*) (see Section 2.1.3).

Consistent with previous data (Watanabe et al. 1998; Galibert et al. 1999), PAX3 activated expression of reporter genes linked to the *MITF* (Fig. 3-13A) and *Trp-1* (Fig. 3-13B) promoter sequences. Surprisingly, our cohort of disease mutations had variable effects on PAX3 regulation of each reporter plasmid. The N47H mutant, which displayed efficient binding to both the *Trp-1* and *MITF* sites, activated both reporter genes to a similar extent as wild type PAX3. In contrast, the G81A mutant and both HD mutants failed to activate *MITF-Luc* and *Trp1-Luc*, and, surprisingly, the G81A and R271G mutants caused significant repression of both reporter constructs below baseline levels. The remaining mutants, *Sp*^{d}, F45L, and S84F, exhibited distinct effects on the two reporters, causing a four- to five-fold increase in *MITF-Luc* expression relative to wild type PAX3, while displaying similar activity to the wild type protein on *Trp1-Luc*. Together, these data indicate that PAX3 disease-causing mutations have distinct effects on gene regulation through the *Trp-1* and *MITF* promoters that do not necessarily correlate with their effects on DNA binding.

3.14 Discussion

A key aspect of understanding PAX3 function requires defining the basis for target gene recognition, a problem that applies to all sequence-specific DNA-binding proteins. Previous analyses of PAX family members have included the derivation of optimal binding sites using *in vitro* selection methods. This has resulted in the identification of high affinity binding sites for the PDs of *Drosophila* Prd, PAX3 (*Nf3'*), PAX2/5/8, and PAX6 (*P6CON*) (Treisman et al. 1991; Czerny et al. 1993; Epstein et al. 1994), the HD of *Drosophila* Prd (*P2*) (Wilson et al. 1993), as well as composite binding sites for the tethered PD and HD of *Drosophila* Prd (*PH0* and *PTE*) (Fujioka et al. 1996; Jun and Desplan 1996). In each case, functional *in vivo* binding sites are expected to resemble these idealized motifs. PAX3 has been shown to regulate the expression of a number of genes, including *c-MET*, *c-RET*, *Myf5*, *MSX2*, *TGFβ2*, and *Dct*, in addition to *MITF* and *Trp-1* (Epstein et al. 1996; Watanabe et al. 1998; Galibert et al. 1999; Lang et al. 2000; Kwang et al. 2002; Lang et al. 2005; Bajard et al. 2006; Mayanil et al. 2006). Although each of these genes' promoters contain motifs similar to the optimized PAX3 PD- and Prd HD-recognition sites, none of them contain perfect consensus sequences, illustrating that *in vitro*-derived, 'optimal' recognition sites may not be representative of all *in vivo* binding sites. Furthermore, we have shown that analyzing the evolutionary conservation of promoter sequences can guide identification of functional *in vivo* target elements. This applies to the *Trp-1* and *c-MET* promoters, which, despite evidence showing that PAX3 can interact with these sequences (Epstein et al. 1996; Galibert et al. 1999; Corry and Underhill 2005a), showed poor conservation among higher eukaryotes.

We have demonstrated that PAX3 employs alternate use of the PD and HD to interact with disparate promoter sequences. In addition, we have shown that individual recognition motifs of composite binding elements need not be absolutely conserved with respect to *in vitro*-derived consensus sequences. For example, the *MITF* and *Mitf* sites contain suboptimal HD motifs, while the *MYF5* and *Myf5* sites possess a PD-binding motif that deviates at one position (see Fig. 3-1). Our results also suggest that interdomain cooperativity mediates high affinity binding to composite sequences composed of suboptimal recognition motifs. Notably, PD binding to the *MITF* sequence is greatly enhanced by the interaction of the HD with an adjacent nonconsensus binding site (Fig. 3-4). In support, the POU$_{HD}$ of OCT family members can interact with AT-rich sequences that deviate from its preferred recognition motif by repositioning itself on the DNA helix, an event that is dependent on adjacent binding of the POU$_S$ domain (Tomilin et al. 2000; Remenyi et al. 2001). Likewise, multimerization of the *Drosophila* bicoid (Bcd) HD supports interaction with distinct binding sites (Dave et al. 2000). Accordingly, the PAX3 HD may also accommodate binding to nonconsensus elements in cases where adjacent PD binding facilitates cooperative binding of the two domains. Together, these observations suggest that cooperative interaction of the PD and HD overcomes the suboptimal nature of the individual recognition motifs.

Inter-domain cooperativity also provides an explanation for the context-dependent effects observed with the PAX3 disease mutants (Fig. 3-11). This was most apparent for the F45L and R271G mutants, which displayed binding to the *MITF* site but were defective in binding to all other sites tested, suggesting that a strict spatial arrangement of the PD and HD allows these mutants to bind to the *MITF* sequence. In contrast, the remaining mutants were either more flexible in accommodating different binding modes (N47H, G81A, and V265F) or consistently displayed inefficient binding (*Sp*^{d} and S84F). Further analyses of the disease mutants revealed additional novel effects on the interaction of PAX3 with target sequences, as shown by the distinct effects of the N47H and G81A PD mutations on both PD binding and cooperativity with the HD. In particular, the N47H mutation enhanced PD binding to the *MITF* sequence as well as formation of a ternary complex with the HD, while the G81A mutation prohibited PD binding and severely limited ternary complex formation (Figs. 3-11B & 3-12A). In addition to establishing that the N47H and G81A mutations alter PD sequence specificity, these results imply that altering the level of cooperativity between the PD and HD contributes to the defects seen in WS. Similarly, while the V265F mutation caused loss of binding to HD-specific sequences (Fig. 3-11B & data not shown), the mutant HD nevertheless maintained a detectable level of cooperativity on the *MITF* sequence (Fig. 3-12B). These observations support the notion that certain PAX3 disease-causing mutations can exert separable effects on the activity of the affected DNA-binding domain and interdomain cooperativity. Importantly, the overall DNA binding competency displayed by the N47H, G81A, and V265F mutants will become relevant to their localization and mobility in cells, as will be discussed in Chapter 4.

Together with the *in vitro* data, our functional analyses reveal that PAX3 can regulate the expression of target genes through the use of alternate modes of DNA binding. Consistent with this idea, PAX3 supported a higher level of activation in the presence of a composite promoter element (*MITF*; 6.9-fold) than on one containing a single PD-binding site (*Trp-1*; 2.7-fold) (Fig. 3-13). Moreover, we noted differences in how these promoters responded to mutant PAX3 proteins. Specifically, the *Sp*^{d}, F45L, and S84F mutants activated transcription several-fold above that of wild type PAX3 only on the *MITF* promoter, while the remaining mutants exhibited similar profiles on both reporter genes. Together, these results indicate that there is no clear correlation between DNA binding affinity and transactivation potential. Similar effects have been observed for disease mutations in the PITX2 transcription factor, where a reduction in DNA binding was nevertheless associated with increased transactivation (Priston et al. 2001). These results further imply that the *Sp*^{d}, F45L, and S84F mutants might cause increased or sustained interactions with transcriptional co-activators or prevent the recruitment of repressive factors, such as GRG4/TLE4 (Lang et al. 2005), to the site in a promoter-dependent manner.

Finally, our luciferase assays reinforce the importance of functional interdependence between the PAX3 PD and HD (Underhill 2000; Apuzzo et al. 2004), since neither HD mutant was able to activate the two reporter constructs (Fig. 3-13), despite the lack of a requirement for the HD in *Trp-1* binding. Additionally, PAX3 G81A repressed transcription of both reporter constructs despite binding to their respective promoter elements with high affinity. Interestingly, the G81A, V265F, and R271G mutants are characterized by significant disparity in *MITF* and *Trp-1* binding (Fig. 3-11C), suggesting they share a functional defect that is independent of the DNA binding mode. In this regard, interactions between the PAX6 PD and the recognition helix of Prd-class HDs support enhanced transactivation of PD-dependent reporters (Mikkola et al. 2001; Bruun et al. 2005), suggesting that inter-domain cooperativity plays an important role in the expression of PAX factor target genes. Together, our examinations of PAX3 disease mutations suggest that variability in promoter-specific effects is likely an important determinant of the WS1 and WS3 phenotypes through alteration of the PAX3 target gene network.

CHAPTER 4. CHARACTERIZATION OF PAX3 IN THE NUCLEUS

Portions of this chapter are published in:

Corry GN, Hendzel MJ, Underhill DA. (2008) Subnuclear localization and mobility of PAX3 are defined by multiple determinants and altered by disease-causing mutations. *Hum. Mol. Gen.*, doi:10.1093/hmg/ddn076

Three-dimensional reconstruction and rendering of nuclei in Figs. 4-2 and 4-4 was performed by Dr. Alan Underhill (University of Alberta, Edmonton, Alberta); FRAP data in Figs. 4-5, 4-8, and 4-9 was collected and compiled by Kristal Missiaen (Cross Cancer Institute, Edmonton, Alberta); PD and HD structures in Fig. 4-10 were created by Dr. Alan Underhill.

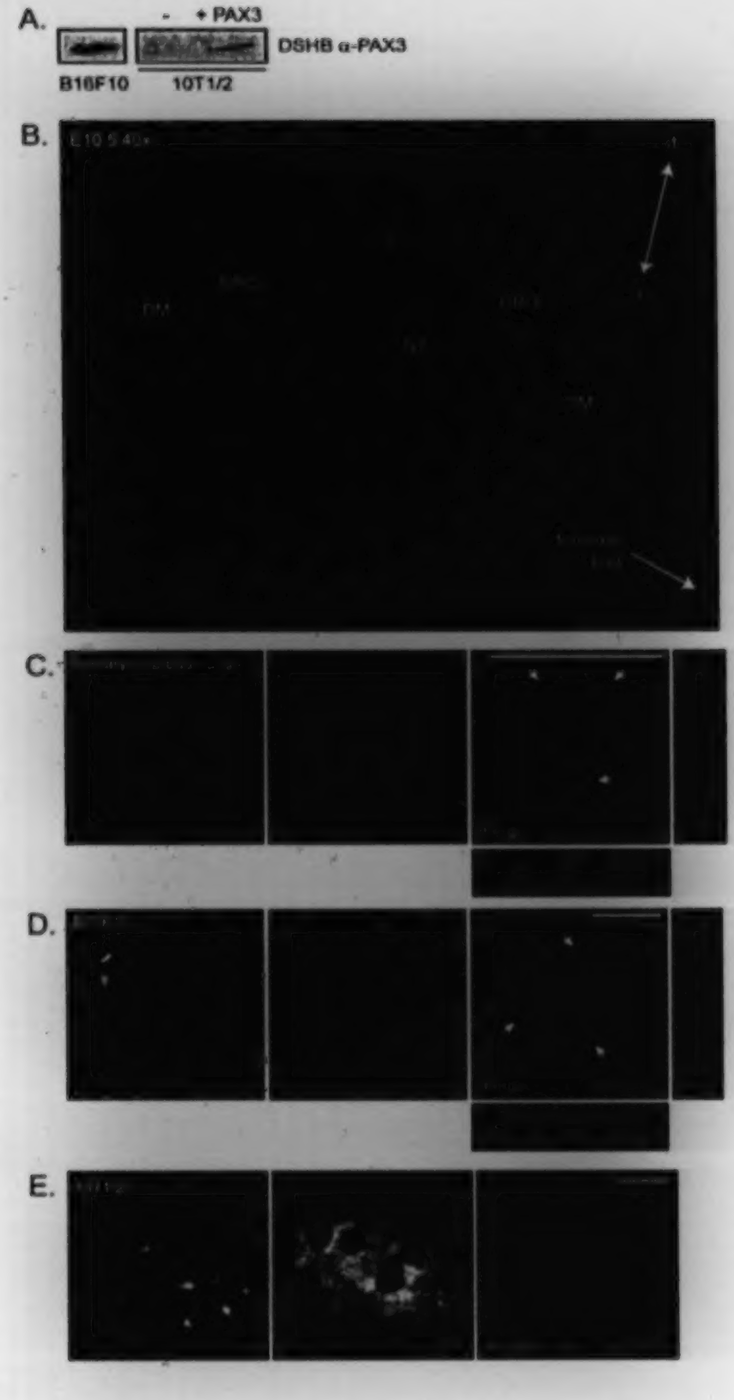
4.1 Background

The eukaryotic nucleus is a highly organized structure and the positioning of nuclear constituents relative to each other plays a vital role in regulating processes such as transcription, DNA replication, and DNA repair (see Section 1.3). Within this scheme, transcription factors exhibit distinct intranuclear distributions and display varying degrees of mobility that facilitates movement between different nuclear compartments in response to signaling events or interactions with other nuclear components. Disease mutations in these transcription factors are thought to contribute in part to pathogenic phenotypes by interfering with subnuclear localization and mobility (see Sections 1.3.3 and 1.3.4 and Corry and Underhill 2005b). Characterization of the *PAX3* locus from WS patients has revealed a broad spectrum of lesions, the majority of which are missense mutations in the PD and HD (Fig. 1-4). Biochemical analyses of these mutants have revealed pleiotropic effects on DNA binding that range from relatively inert to greatly reduced or elevated (Chalepakis et al. 1994a; Epstein et al. 1996; Fortin et al. 1997; Watanabe et al. 1998; Corry and Underhill 2005a; see Sections 3.11 and 3.12). In addition, reporter gene assays using the melanocyte-specific PAX3 targets *MITF* (Watanabe et al. 1998) and *Trp-1* (Galibert et al. 1999) indicate that disease mutations have diverse effects on promoter activity (Corry and Underhill 2005a). Based on these assays, it has not been straightforward to understand the molecular basis of PAX3 loss of function or establish relationships between phenotype and genotype, particularly in accounting for the differences in WS types 1 and 3. Complicating matters, the same *PAX3* mutant allele can produce distinct phenotypes, even within the same family. Consequently, it appears that other aspects of PAX3 function may be compromised by disease-causing mutations, potentially accounting for the variable phenotypic expressivity seen in WS.

With this in mind, we have analyzed the intranuclear behavior of PAX3 and disease-associated mutants. Immunofluorescence analyses indicate that the majority of PAX3 occupies the interchromatin compartment, with only sporadic co-localization with sites of transcription. Moreover, this distribution is retained by both untagged and GFP-fused PAX3, providing a means to characterize PAX3 disease alleles. Consistent with previous studies where the PD was shown to modulate DNA binding by the HD (Underhill et al. 1995; Fortin et al. 1997; Underhill and Gros 1997), functional interaction between the two domains also influences the steady state distribution of PAX3 and suggests the presence of multiple determinants for the subnuclear localization of PAX3. Upon assessment of their intranuclear localization and dynamics, PAX3 mutants fell into two distinct mechanistic categories regardless of the structural proximity of the mutated residues. The first group (class I) exhibits a diffuse distribution and markedly increased mobility when compared to wild type PAX3, while class II mutants display evidence of subnuclear compartmentalization and mobility intermediate between wild type PAX3 and class I proteins. Importantly, our results indicate that DNA binding is not a primary determinant of PAX3 distribution and movement. Together, these results establish that altered localization and dynamics play a key role in PAX3 dysfunction and that loss of the underlying determinants represents the principal defect for a subset of WS mutations.

4.2 Subnuclear localization of PAX3

To investigate the intranuclear distribution of endogenous PAX3, we performed immunofluorescence on whole embryo sections, primary limb bud cultures, and B16F10 mouse melanoma cells using an antibody that specifically detects endogenous PAX3 in B16F10 whole cell extracts and exogenously expressed PAX3 in 10T1/2 mouse embryonic fibroblast extracts (Fig. 4-1A). In 10.5 dpc embryos, PAX3 immunofluorescence was excluded from Hoechst-stained regions over the entire PAX3 expression domain, including cells of the dorsal neural tube, dorsal root ganglia, dermomyotome, and migrating myoblasts (Fig. 4-1B). To investigate the subnuclear distribution of PAX3 in more detail, we performed immunofluorescence on endogenous PAX3 in B16F10 cells and primary limb bud cultures, as well as in mouse 10T1/2 cells after transient transfection with PAX3 expression plasmids. An important objective of these analyses was to establish a reference point for the characterization of PAX3 disease alleles. Primary limb bud culture nuclei typically displayed a reticular distribution of PAX3 showing no overlap with the Hoechst signal (Fig. 4-1C). In these cells, a number of bright foci were often observed, with a large quantity of these in the nuclear periphery (*arrows*). In B16F10 cells PAX3 adopted a similar punctate pattern, showing no overlap with pericentromeric or perinuclear heterochromatin or nucleoli (Fig. 4-1D). As in the primary cultures, we observed elevated intensity of several PAX3 foci per nucleus (*arrows*), although PAX3 appeared to be excluded from perinuclear regions in B16F10 cells. In 10T1/2 fibroblast cells transiently transfected with a plasmid expressing untagged PAX3, we observed a similar intranuclear distribution to that of endogenous PAX3 (Fig. 4-1E). Interestingly, the PAX3 signal intensity seemed increased in


A.
- + PAX3
DSHB α-PAX3
B16F10
10T1/2
B.
E10.5 40x
DM
NT
DM
C.
D.
E.

F.

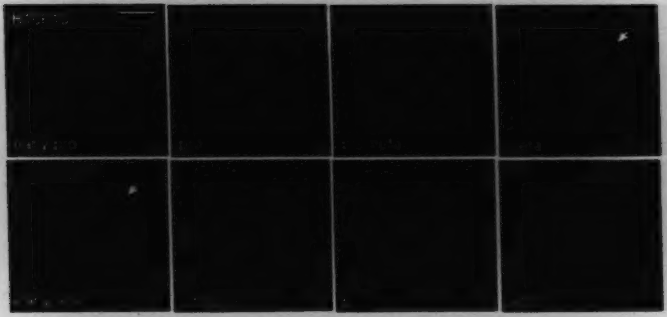


Figure 4-1. Subnuclear localization of PAX3. (A) Western blot of endogenous PAX3 in B16F10 (*left*) and 10T1/2 cells, either untransfected (*right*, *lane 1*) or transfected with a PAX3 expression plasmid (*right*, *lane 2*). In both experiments, the DSHB anti-PAX3 antibody was used (see Section 2.9). Indirect immunofluorescence was used to examine the subnuclear localization of endogenous PAX3 in (B) 10.5 dpc mouse embryos (DM, dermomyotome; DRG, dorsal root ganglia; NT, neural tube; L, lumen; dorsal-ventral axis is indicated), (C) 11.5 dpc primary limb bud cultures, and (D) B16F10 murine melanoma cells. Cells were treated with the DSHB PAX3 antibody (green) and co-stained with Hoechst 33258 (red). Merged panels in (C) and (D) include X-Y (*main*), X-Z (*bottom*), and Y-Z (*side*) projections; arrows in (C) and (D) indicate bright foci characteristic of endogenous PAX3. (E) Expression of untagged PAX3 in 10T1/2 fibroblasts. Grayscale panels are unaltered; merged panels have been deconvolved. (F) Indirect immunofluorescence of unsynchronized B16F10 cells using the DSHB PAX3 antibody was used to analyze spatial and temporal distribution of PAX3 during mitotic stages. Cells in different stages of mitosis were identified by staining with Hoechst. Panels are comprised of merged, deconvolved Hoechst (red) and PAX3 (green) signals. Bar, 10 µm.

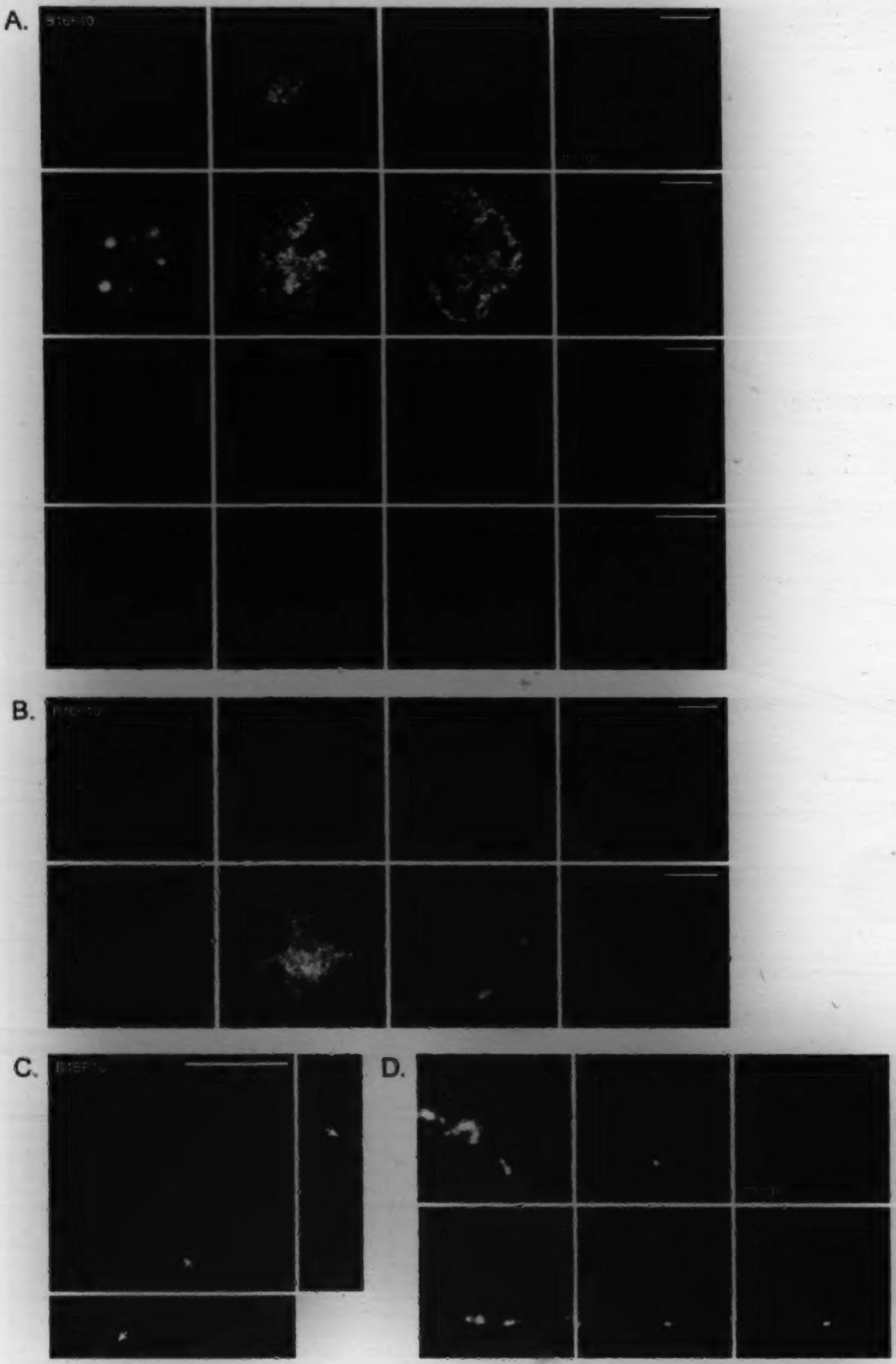
perinucleolar areas in 10T1/2 cells and PAX3 appeared more diffuse in 10T1/2 cells than in B16F10 cells, likely due to overexpression. C2C12 mouse myoblasts transfected with a PAX3 expression plasmid showed a similar intranuclear pattern to that seen in 10T1/2 cells (not shown).

We also monitored PAX3 localization throughout mitosis in unsynchronized B16F10 cells stained with Hoechst (Fig. 4-1F). In prophase, PAX3 was initially confined to the nucleus at early stages and eventually redistributed to the cytoplasm. This pattern was maintained until cells entered telophase and the chromosomes began to decondense, at which point PAX3 appeared in proximity to loosely packed chromatin. By late telophase, the majority of PAX3 was found in nuclei. As observed during interphase, PAX3 was clearly excluded from regions of Hoechst staining, which was most apparent with highly condensed chromosomes at metaphase. Despite the redistribution of PAX3 during mitosis, the protein retained its reticular pattern observed during interphase where PAX3 immunofluorescence often appears as 'strands' (*arrows* in Fig. 4-1F). Together, these analyses reveal that the steady state distribution of PAX3, both *in vivo* and *in situ*, involves a mesh-like pattern during interphase (and to a lesser extent in mitosis) and shows no co-localization with heterochromatic landmarks demarcated by Hoechst staining at any point during the cell cycle.

4.3 Co-localization of PAX3 and post-translational histone modifications

The DNA binding dyes Hoechst 33258 and 4',6-diamidino-2-phenylindole (DAPI) exhibit a preference for AT-rich DNA, which is more abundant in heterochromatin and as a result, do not efficiently demarcate the euchromatic compartment. To further assess the subnuclear localization of PAX3, indirect immunofluorescence was used to visualize active and inactive chromatin using antibodies that recognize histone modifications enriched in these chromatin domains. AcH3 and AcH4 mark transcriptionally competent euchromatin (Eberharter and Becker 2002), H3K4me3 marks the promoter region of transcriptionally active genes (Santos-Rosa et al. 2002; Ng et al. 2003), and H3K36me3 is associated with transcriptional elongation by RNAPII (Krogan et al. 2003; Li et al. 2003; Xiao et al. 2003). Each of these euchromatic chromatin marks displayed a reticular pattern that was excluded from Hoechst, although more diffuse than PAX3 (Fig. 4-2A). Modifications associated with heterochromatin and silenced chromatin were also investigated, including H3K9me3, found primarily in condensed, silenced heterochromatin (Lachner et al. 2001; Peters et al. 2002) and H4K20me3, a marker of pericentromeric heterochromatin (Schotta et al. 2004). Not surprisingly, PAX3 failed to co-localize with these heterochromatin marks, given that they replicate the Hoechst pattern (Fig. 4-2B).

As the euchromatin-associated histone modifications revealed qualitatively similar localization with respect to each other, H3K4me3 was used to demonstrate their relationship with PAX3. PAX3 immunofluorescence was in close proximity to H3K4me3 throughout the nucleus, where the signals appeared interdigitated or intertwined and also intersected at some points (*arrows* in Fig. 4-2C). This is more obvious upon introducing an intensity threshold (Fig. 4-2D) and reveals that PAX3 is closely apposed to euchromatin with a small subset displaying co-localization. Lastly, rendering images in 3D illustrates the clear difference in localization of PAX3 with respect to euchromatin and heterochromatin (Fig. 4-2E). Based on this spatial


A.
B16F10
merge
B.
B16F10
C.
B16F10
D.
merge

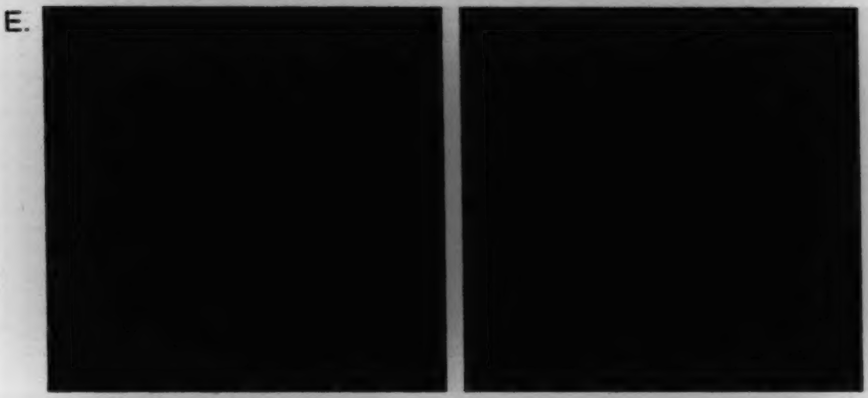


Figure 4-2. Distribution of PAX3 relative to post-translational histone modifications. (A) Co-localization between PAX3 and histone modifications associated with euchromatin, including AcH3 (*top row*), AcH4 (*second row*), H3K4me3 (*third row*), or H3K36me3 (*bottom row*), were analyzed in B16F10 cells by indirect immunofluorescence. (B) Co-localization between PAX3 and histone modifications characteristic of transcriptionally silent heterochromatin, including H3K9me3 (*top row*) and H4K20me3 (*bottom row*) in B16F10 cells. Grayscale panels are unaltered; merged panels have been deconvolved. Bar, 10 µm. (C) Three-dimensional rendering of PAX3 and H3K4me3 in B16F10 cells. X-Y (*main*), X-Z (*bottom*), and Y-Z (*side*) projections are shown; arrows indicate sites of overlap between PAX3 and H3K4me3. (D) Detail of co-localization between PAX3 and H3K4me3 in B16F10 cells. Top and bottom rows represent unique regions of nucleus in (C) with thresholds to illustrate their spatial proximity. (E) Three-dimensional rendering of segments from nuclei shown in Fig. 4-1D (*left*) and Fig. 4-2C (*right*) detailing co-localization between endogenous PAX3 and Hoechst-stained chromatin and H3K4me3, respectively.

relationship, we propose that the vast majority of PAX3 occupies the interchromatin space and that only a subset appears to be engaged with euchromatin.

4.4 Co-localization of PAX3 and sites of transcriptional activity

The spatial relationship of PAX3 with transcriptionally active chromatin was further evaluated by incubating cells with 5-FUrd, a nucleoside analog that is incorporated into newly synthesized RNA and can be recognized by a BrdU-specific antibody. We incubated B16F10 cells or 10T1/2 cells transfected with a PAX3 expression plasmid with 5-FUrd for 30 min, then performed co-immunofluorescence with PAX3- and BrdU-specific antibodies. In B16F10 cells, the BrdU antibody recognized numerous small punctate nuclear foci corresponding to sites of 5-FUrd incorporation (Fig. 4-3, *top row*) and the PAX3 signal showed partial overlap with several 5-FUrd foci per nucleus (*inset*). A similar pattern of co-localization between PAX3 and 5-FUrd was observed in 10T1/2 nuclei (Fig. 4-3, *bottom row*), although the distribution of 5-FUrd incorporation did not resemble that seen in B16F10 cells. Consistent with the juxtaposition of PAX3 and H3K4me3, these results provide evidence that PAX3 is found in regions of transcriptional activity, but suggest that only a small number of PAX3 sites contain transcriptionally active loci at any given time.

4.5 PAX3 DNA-binding domains constitute separable determinants of subnuclear localization

To examine intrinsic determinants of PAX3 subnuclear compartmentalization, we created a series of GFP fusions of the full-length protein and DNA-binding domains (Fig.

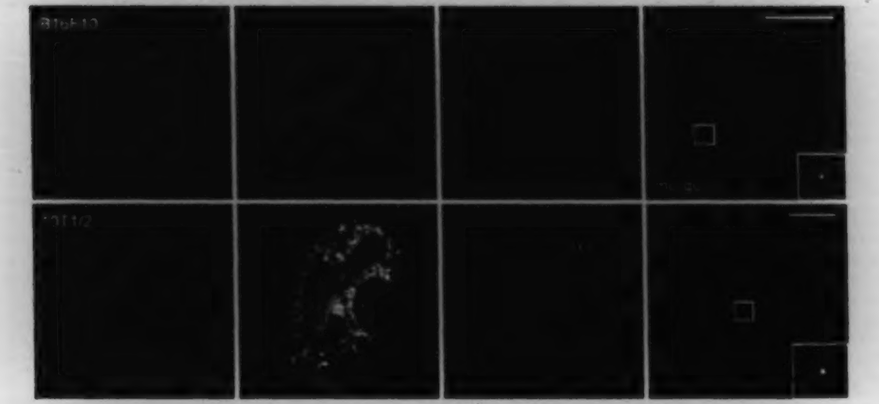


Figure 4-3. PAX3 shows limited overlap with sites of transcriptional activity. Co-localization between endogenous (*top*) or transiently transfected (*bottom*) PAX3 and sites of RNA synthesis. Cells were grown for 24 h with (10T1/2) or without (B16F10) transfection, incubated with 1 mM 5-fluorouridine (5-FUrd) at 37°C for 30 min, then examined by indirect immunofluorescence using antibodies specific for PAX3 and bromodeoxyuridine. Inset, detail of co-localization between PAX3 and 5-FUrd foci (yellow). Grayscale panels are unaltered; merged panels have been deconvolved. Bar, 10 µm.

4-4A). Initial investigations were performed with full length PAX3, tagged at either the amino- (GFP-PAX3) or carboxy-terminus (PAX3-GFP) and the PD or HD, either individually (PD-GFP and HD-GFP) or in the same molecule with (N-PDHD-GFP) and without (PDHD-GFP) the amino-terminal extension. PAX3-GFP displayed a similar distribution as untagged PAX3 expressed in 10T1/2 cells (Figs. 4-4B, *left* & 4-4C, *left, top row*). This pattern was observed in approximately 80% of transfected cells, while in the remaining cells, PAX3-GFP co-localized to a large extent with heterochromatic areas (Fig. 4-4B, *right*). To ensure that the GFP moiety was not causing this distribution by affecting PAX3 conformation or stability in the nucleus, we expressed PAX3-GFP or GFP-PAX3 in 10T1/2 cells and performed immunofluorescence with the DSHB anti-PAX3 antibody. The antibody detected each PAX3 fusion protein, and the GFP and DSHB signals were virtually superimposable (not shown), demonstrating that fusing GFP to either terminus of PAX3 does not hinder folding or stability in the nucleus. PAX3 tagged with GFP at the amino terminus showed the same distribution as PAX3-GFP (not shown). The N-PDHD-GFP protein exhibited two distinct intranuclear distributions: one that featured exclusion from Hoechst-stained regions (Fig. 4-4C, *left, middle row*) and one that displayed considerable co-localization with Hoechst-rich areas (Fig. 4-4C, *left, bottom row*). PDHD-GFP produced a pattern that resembled PAX3-GFP (Fig. 4-4C, *right, top row*), but unlike N-PDHD-GFP, showed no evidence of localization to chromatin, suggesting the amino-terminal region of PAX3 may contain an element that promotes interaction with condensed chromatin. PD-GFP formed a mesh-like pattern that appeared less compartmentalized than full length PAX3-GFP or PDHD-GFP, but remained excluded from heterochromatin (Fig. 4-4C, *right, middle row*). Moderate

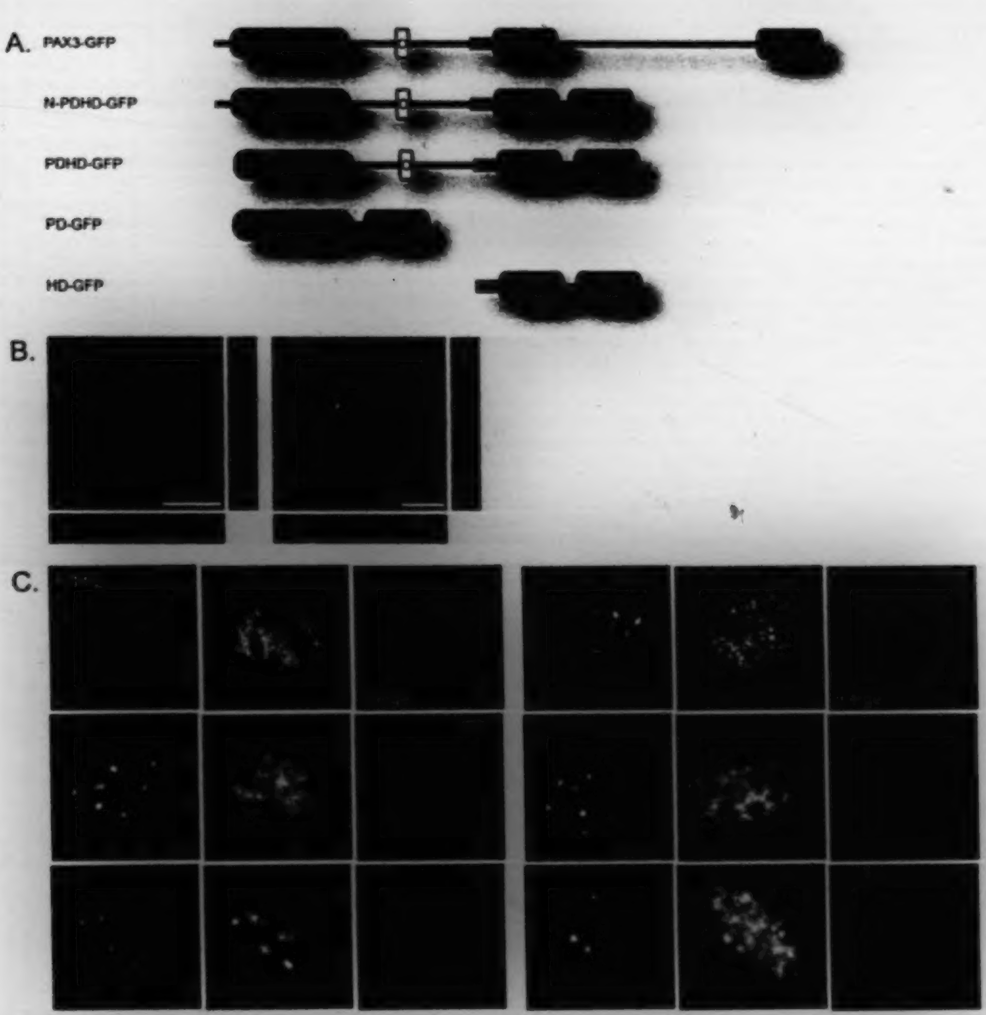


Figure 4-4. The PAX3 paired domain and homeodomain represent separable determinants of intranuclear localization. Intrinsic determinants of PAX3 were analyzed by fusing green fluorescent protein (GFP) to the full length protein or the PAX3 DNA-binding domains. (A) GFP-PAX3 fusions (not to scale) used to address subnuclear compartmentalization. Locations of the paired domain (PD; red), octapeptide (O; yellow), homeodomain (HD; blue), nuclear localization signal (NLS), and GFP are shown. (B) Subnuclear localization of full-length PAX3-GFP. In approximately 20% of 10T1/2 cells transfected with a PAX3-GFP expression plasmid, localization of the protein to densely staining chromatin domains was observed (*right*). Panels show three-dimensional rendering of PAX3-GFP expressed in 10T1/2 cells comparing non-heterochromatic (*left*) and heterochromatic (*right*) localization. Panels show X-Y (*main*), X-Z (*bottom*), and Y-Z (*side*) projections. (C) Subnuclear localization of PAX3-GFP (*left, top row*), N-HD-GFP (*left, second & third rows*), PDHD-GFP (*right, top row*), PD-GFP (*right, middle row*) and HD-GFP (*right, bottom row*) in 10T1/2 cells. Grayscale panels are unaltered; merged panels have been deconvolved. Bar, 10 μm.

cytoplasmic staining was also observed and may be due to the lack of a canonical NLS. Unexpectedly, HD-GFP consistently displayed extensive co-localization with Hoechst-stained chromatin (Fig. 4-4C, *right, bottom row*), including pericentromeric heterochromatin. Together, these results imply that PAX3 contains determinants for localization to both the interchromatin compartment and condensed chromatin, but that the former predominates in the intact protein.

In addition to subcellular compartmentalization, transcription factor mobility has emerged as an essential component of gene regulation. We therefore used fluorescence recovery after photobleaching (FRAP) experiments to analyze the mobility of PAX3-GFP and the isolated GFP-tagged DNA-binding domains, either tethered (PDHD-GFP) or in isolation (PD-GFP, HD-GFP). Individually, PD-GFP and HD-GFP exhibited greatly increased mobility compared to PAX3-GFP (Fig. 4-5A), while PDHD-GFP was considerably less mobile relative to the full length protein. In a FRAP recovery curve, the value at the first time point after photobleaching (1.5 s) provides an approximate measurement of the protein's freely diffusing pool. Us[illegible] this point of reference, our results show that PD-GFP possesses the largest pool of freely diffusing protein, followed by HD-GFP, PAX3-GFP, and PDHD-GFP (Fig. 4-5B). Although HD-GFP displayed a lower effective recovery rate than the PD, it was the first to recover to 100% (after only 20 s) and was the only protein to fully recover after 90 s, suggesting that a fraction of PD-GFP, PDHD-GFP, and PAX3-GFP is immobile or exhibits reduced mobility. Importantly, the discrepancy between the recovery of the isolated DNA-binding domains and full length PAX3 indicates that DNA binding is not the primary determinant of PAX3 mobility in the nucleus. Furthermore, the slower recovery rate of PDHD-GFP

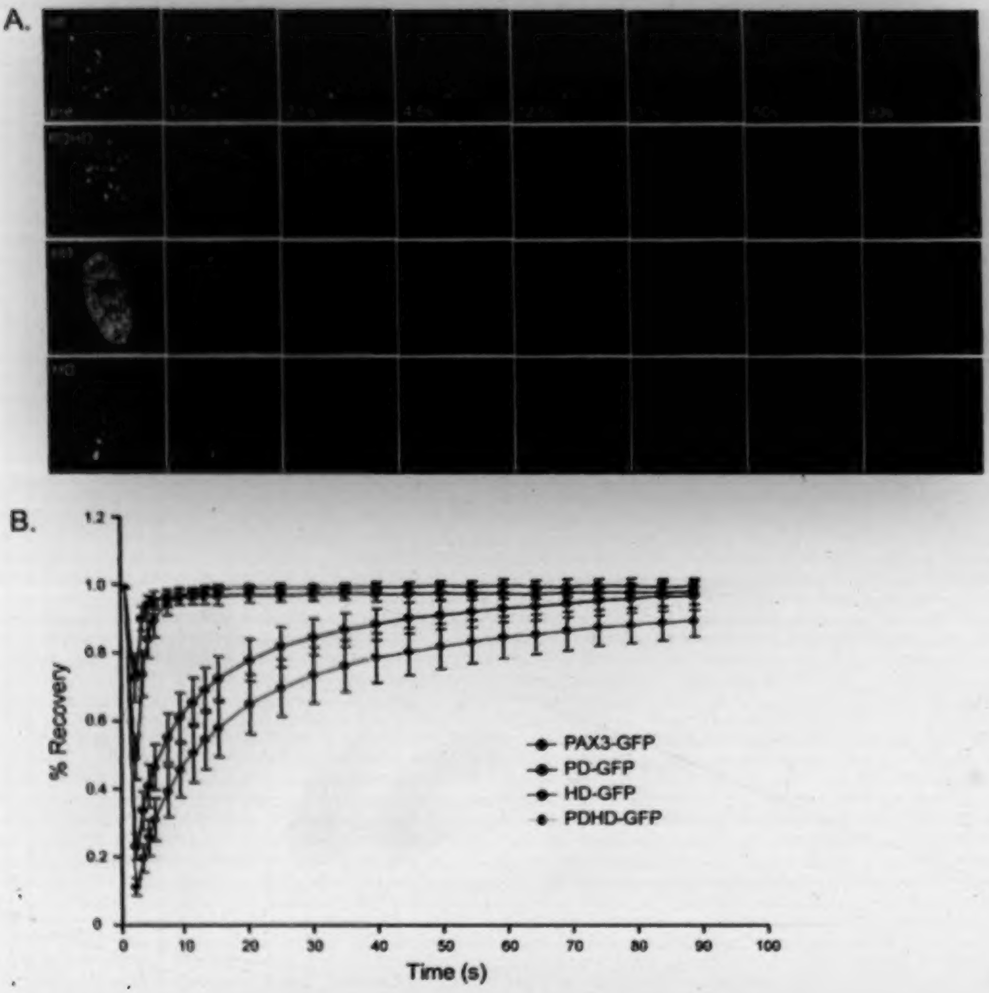


Figure 4-5. Tethering the PAX3 DNA-binding domains constrains their mobility in live cells. Intranuclear mobility of full length PAX3 and its DNA-binding domains was examined using fluorescence recovery after photobleaching (FRAP). (A) Intracellular localization of PAX3-GFP, PD-GFP, HD-GFP, and PDHD-GFP in live cells. (B) FRAP recovery curves of PAX3-GFP, PD-GFP, HD-GFP, and PDHD-GFP in mouse 10T1/2 cells. Plots show recovery versus time; each plot represents the average of thirty individual FRAP experiments. Standard deviation at each time point is indicated by error bars.

compared to PAX3-GFP suggests that regions not encompassed by the PDHD protein influence PAX3 behavior in the nucleus.

4.6 Effect of flanking regions on PAX3 homeodomain subnuclear localization

The extensive co-localization between HD-GFP and Hoechst-stained regions was unexpected and suggests that the PAX3 HD may be a determinant for interaction with heterochromatin, albeit one that is suppressed in the context of the full length protein. Tethering the PD and linker to the HD precluded localization to condensed chromatin (Fig. 4-4C); however, these results did not specify the location of the determinants that suppress HD-chromatin interaction. To investigate this, we fused portions of the interdomain linker and carboxy-terminal region of PAX3 to the HD (Fig. 4-6A) and analyzed localization of these proteins in 10T1/2 cells. Fusion of the PAX3 carboxy-terminal region distributed the HD away from Hoechst-rich regions both in the presence (Fig. 4-6B, *top row*) and absence (Fig. 4-6B, *second row*) of the linker. Interestingly, the intranuclear pattern of HD-C-GFP bore a striking resemblance to GFP-PAX3 (*cf.* Fig. 4-4C), given its punctate distribution, increased levels in perinucleolar regions, and exclusion from the nuclear periphery. These observations imply that the carboxy-terminus of PAX3 harbors essential components, possibly the transactivation domain, that direct PAX3 to its appropriate nuclear locale. These data are also consistent with results demonstrating that the PAX3 transactivation domain can regulate the activity of the HD (Cao and Wang 2000). Addition of the full linker slightly altered subnuclear localization of the HD, as the linker-HD-GFP protein displayed only partial co-localization with pericentromeric heterochromatin (Fig. 4-6B, *third row*), while deletion

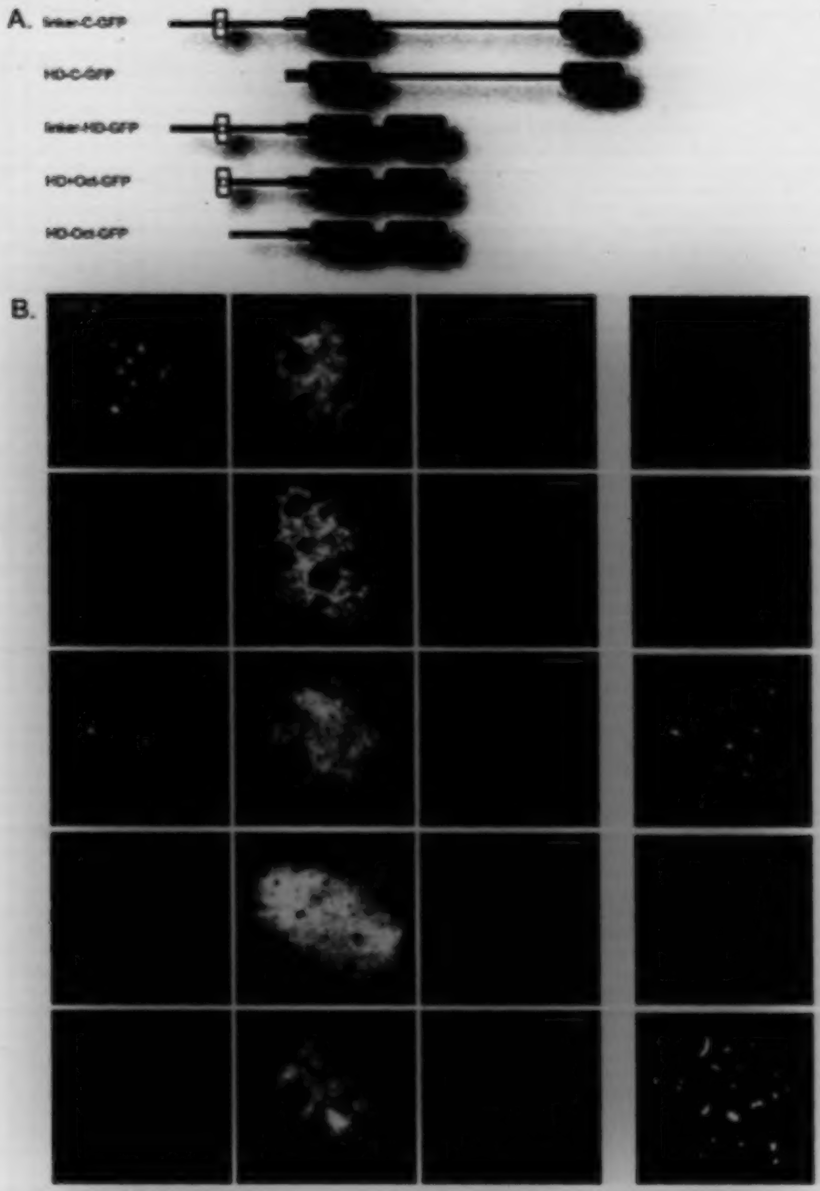


Figure 4-6. Flanking regions influence the subnuclear localization of the PAX3 homeodomain. (A) Schematic representation of GFP fusions (not to scale) used to analyze effects of flanking regions on HD distribution in the nucleus. HD, homeodomain; O, octapeptide; NLS, nuclear localization signal. (B) Subnuclear localization of the linker-C-GFP (*top*), HD-C-GFP (*second row*), linker-HD-GFP (*third row*), HD+Oct-GFP (*fourth row*), and HD-Oct-GFP (*bottom row*) proteins in 10T1/2 cells. Signal intensity thresholds (*right column*) detail regions of overlap (yellow) between Hoechst (red) and GFP (green). Grayscale panels are unaltered; merged panels have been deconvolved. Bar, 10 μm.

of linker sequence amino-terminal to the octapeptide caused exclusion from compacted chromatin (Fig. 4-6B, *fourth row*). Surprisingly, deletion of the amino-terminal linker and octapeptide caused a reassociation with pericentromeric heterochromatin (Fig. 4-6B, *bottom row*). Together, these results suggest that regions flanking the HD possess important determinants of intranuclear compartmentalization that may influence HD activity in the nucleus. However, the manner by which these determinants regulate the HD is obviously complex, as indicated by the variable localization patterns upon deletion into the interdomain linker.

4.7 Effects of disease mutations on PAX3 subnuclear localization

Our observations establish that PAX3-GFP recapitulates the distribution of untagged and endogenous PAX3, and provides a means to characterize PAX3 mutants independent of effects the mutations may have on antibody recognition. For this, we used the same cohort of PD and HD mutations previously shown to exert variable effects on DNA binding and reporter gene transactivation (Corry and Underhill 2005a; see Sections 3.11-3.13). We also assessed the effects of two WS-associated mutations outside the DNA-binding domains, A196T in the linker and Q391H in the carboxy-terminus (see Fig. 4-7A for positions of mutations). To compare the localization of wild type and mutant PAX3 proteins, each was expressed as a GFP fusion protein in 10T1/2 cells. As noted for wild type PAX3-GFP (Fig. 4-7B), all mutants were excluded from Hoechst-stained chromatin (Fig. 4-7C-E). The *Sp*^{d} mutant displayed a more uniform distribution than wild type PAX3-GFP (Fig. 4-7C, *top row*), was less abundant in perinucleolar areas, and tended to display one or two large bright foci in some cells rather

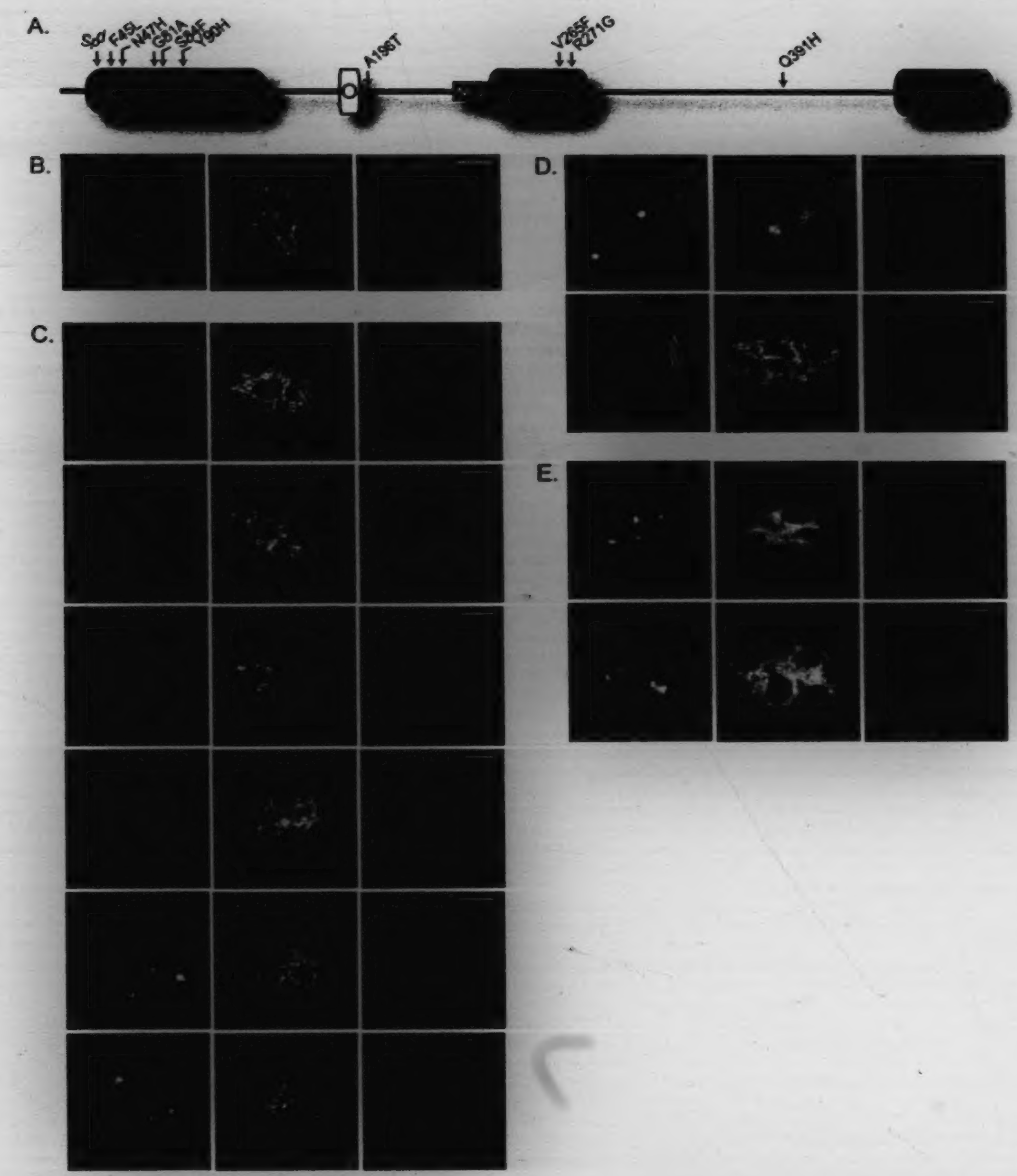

A.
Sp^d
F45L
N47H
G81A
S84F
Y90H
A196T
V265F
R271G
Q391H
B.
C.
D.
E.

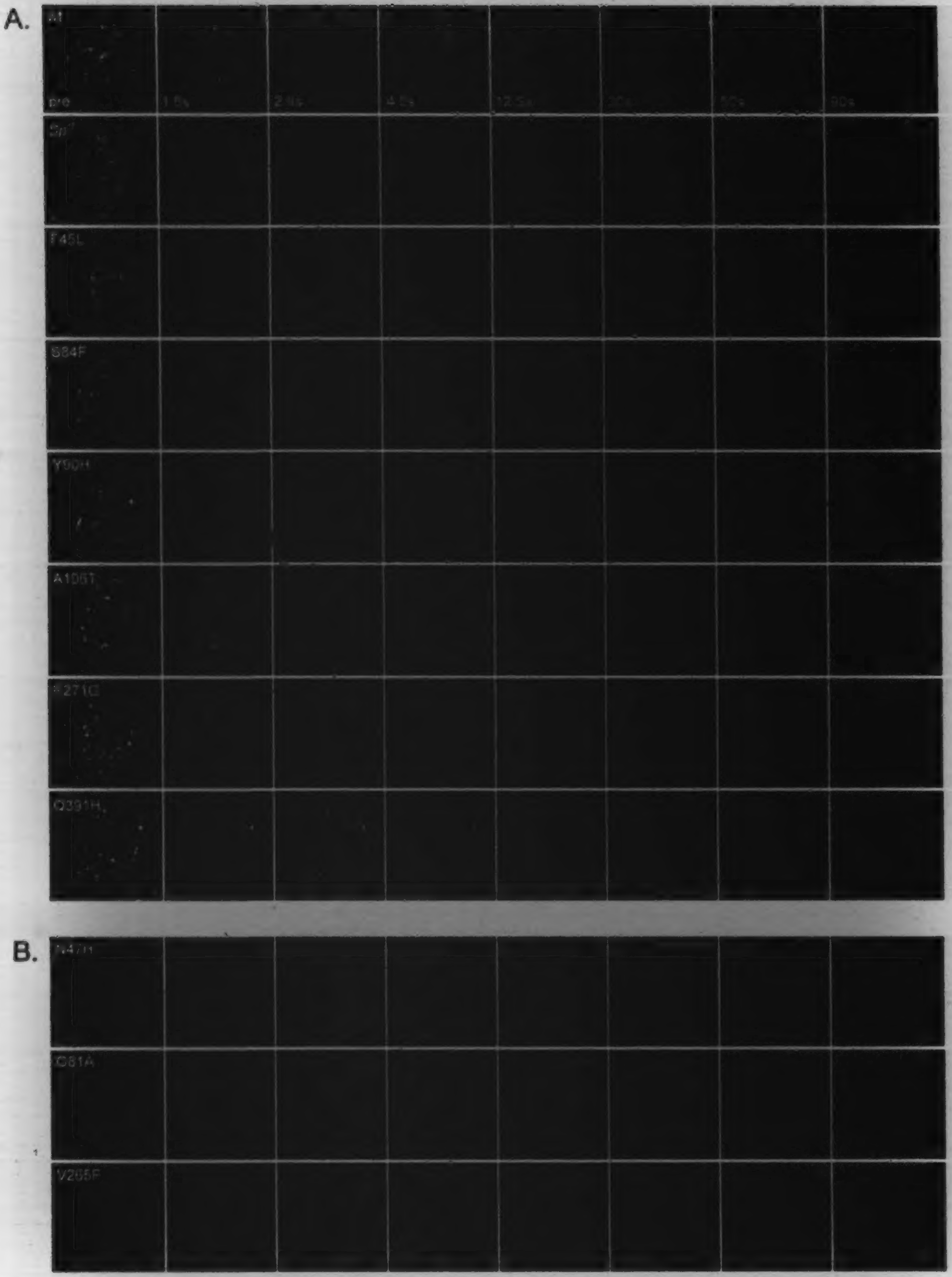
Figure 4-7. Disease mutations alter the subnuclear localization of PAX3. (A) Schematic of PAX3 (not to scale) showing locations of disease-causing mutations analyzed for effects on distribution in the nucleus. Subnuclear compartmentalization of wild type PAX3-GFP (B) and variants of PAX3-GFP containing mutations in the PD (C), HD (D), linker (E; *top row*), or carboxy-terminus (E; *bottom row*) is shown. Grayscale panels are unaltered; merged panels have been deconvolved. Bar, 10 µm.

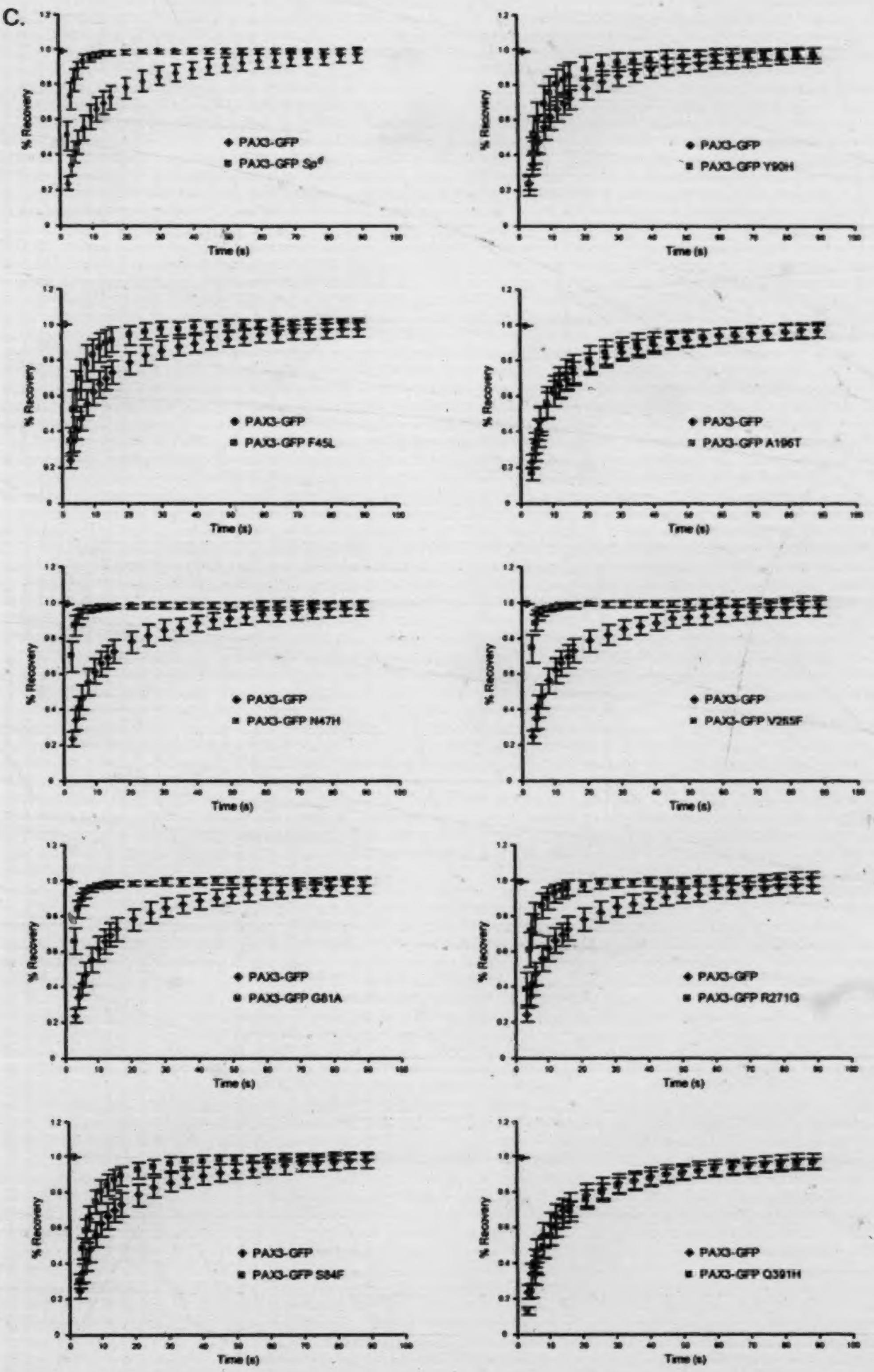
than the multiple bright foci characteristic of PAX3-GFP. The F45L mutant displayed a similar pattern to the wild type protein, with a number of bright foci and concentrations around nucleoli (Fig. 4-7C, *second row*). Both the N47H (Fig. 4-7C, *third row*) andG81A (Fig. 4-7B, *fourth row*) proteins showed consistently lower expression levels than the other PAX3 variants, tended to be excluded from perinucleolar areas, and generally showed a more diffuse staining pattern than the wild type protein. These mutants also showed a moderate level of cytoplasmic localization, suggesting the N47H and G81A mutations may interfere with nuclear localization or retention of PAX3. In addition, the N47H mutant consistently showed exclusion from the nuclear periphery. The S84F mutant was similar in appearance to wild type and F45L (Fig. 4-7B, *fifth row*), although it tended to exhibit a noticeably more diffuse background localization pattern than either. Likewise, the Y90H mutant displayed numerous bright foci against a less intense, diffuse background (Fig. 4-7C, *bottom row*). The HD mutant V265F was distributed throughout the nucleus and cytoplasm (Fig. 4-7D, *top row*), closely resembling the N47H and G81A mutants. Finally, the R271G HD variant showed a pattern similar to *Sp*^{d}, having a uniform distribution in Hoechst-free areas with less enrichment around the nucleoli and one or more large bright foci (Fig. 4-7D, *bottom row*). The A196T and Q391H mutants showed similar characteristics to wild type PAX3-GFP, including a punctate pattern and enrichment in perinucleolar areas (Fig. 4-7E), although we did not observe the same bright foci distinctive of PAX3-GFP. In this regard, the A196T and Q391H mutants more closely resembled the *Sp*^{d} and R271G patterns. Together, these observations demonstrate that PAX3 disease mutations influence subnuclear compartmentalization to differing degrees, which may represent another component of PAX3 dysfunction.

4.8 Nuclear dynamics of wild type and mutant PAX3

Having established a baseline for PAX3-GFP dynamics (Fig. 4-5), we next investigated the intranuclear mobility of our cohort of PAX3 disease mutants using FRAP. At the same time, this provided an opportunity to compare the localization of wild type PAX3-GFP to disease-causing mutants in live and fixed cells. The post-bleach recovery profiles of wild type and mutant PAX3-GFP proteins show that their localization in live cells can be generally correlated with that observed in fixed cells and allow categorization of the mutants into two groups. The first includes Sp^d, F45L, S84F, Y90H, and R271G, which retain a compartmentalized appearance similar to that of wild type PAX3 (Fig. 4-8A). The N47H, G81A, and V265F mutants form a second group characterized by diffuse distribution and, as noted in fixed cells (Fig. 4-7), display varying degrees of cytoplasmic localization, although fluorescence intensity in the nucleus is greater in each case (Fig. 4-8B). As will be elaborated below, the presence or absence of a compartmentalized distribution correlates with mobility.

The analysis of PAX3 disease allele dynamics provides two major findings: with the exception of PAX3-GFP Q391H, all mutants tested displayed increased mobility when compared to the wild type protein and had essentially recovered to 100% by 90s (Fig. 4-8C), indicating they lack the immobile fraction seen with wild type PAX3. Nevertheless, the mutants fell into two categories that could be distinguished in terms of overall mobility and free pool sizes. The first group, designated class I, includes the


A.
wt
pre
1.5s
2.5s
4.5s
12.5s
20s
50s
90s
Sn
F45L
S84F
Y90H
A106T
R271Q
Q391H
B.
N47H
C81A
V265F


C.
% Recovery
Time (s)
0
0.2
0.4
0.6
0.8
1.0
1.2
0
10
20
30
40
50
60
70
80
90
100
◆ PAX3-GFP
■ PAX3-GFP Sp^{d}
◆ PAX3-GFP
■ PAX3-GFP Y90H
◆ PAX3-GFP
■ PAX3-GFP F45L
◆ PAX3-GFP
■ PAX3-GFP A196T
◆ PAX3-GFP
■ PAX3-GFP N47H
◆ PAX3-GFP
■ PAX3-GFP V265F
◆ PAX3-GFP
■ PAX3-GFP G81A
◆ PAX3-GFP
■ PAX3-GFP R271G
◆ PAX3-GFP
■ PAX3-GFP S84F
◆ PAX3-GFP
■ PAX3-GFP Q391H

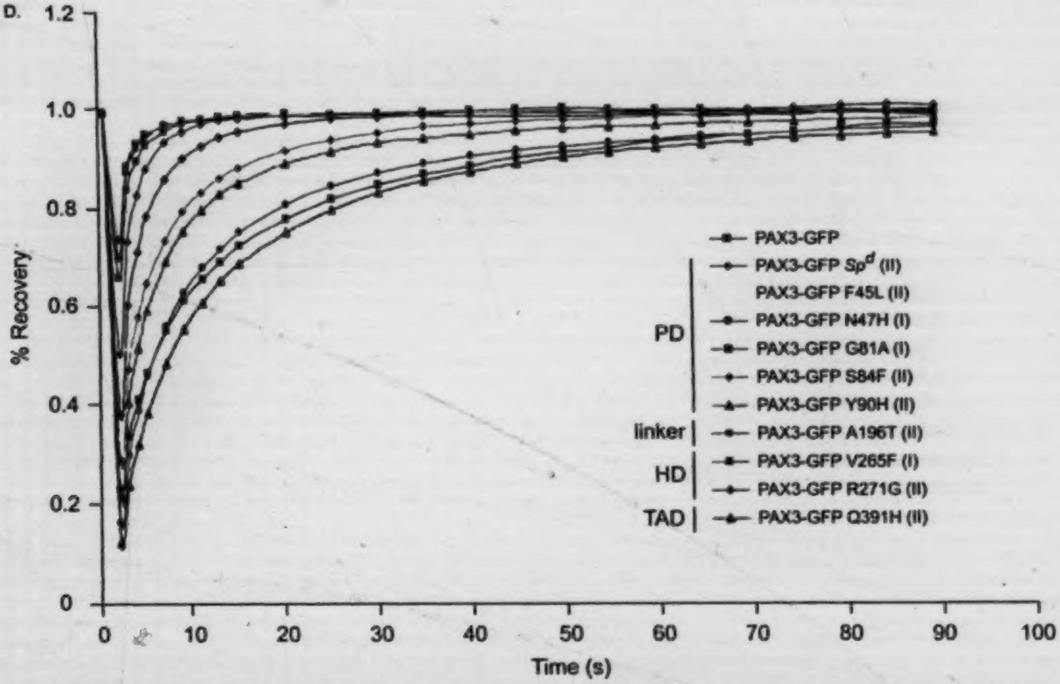


Figure 4-8. Behavior of wild type and mutant PAX3-GFP in live cells. Intranuclear mobility of full-length wild type and mutant PAX3-GFP was analyzed using FRAP in transiently transfected 10T1/2 cells. Proteins that exhibit similar recoveries, (A; class II) wt, *Sp*^{d}, F45L, S84F, Y90H, A296T, R271G, and Q391H, and (B; class I) N47H, G81A, and V265F, are grouped together. (C) FRAP recovery plots for each mutant PAX3-GFP protein (pink squares) are shown with the wild type PAX3-GFP recovery curve (blue diamonds) for comparison. Plots show recovery versus time; each plot represents the average of thirty individual FRAP experiments. Error bars indicate standard deviation at each time point. (D) FRAP recovery curves for wild type and mutant PAX3-GFP proteins. The location of each mutation (PD, paired domain; HD, homeodomain; TAD, transactivation domain) and the mutational class (I or II) each mutant belongs to is indicated. Recovery curves for wild type and mutant PAX3-GFP are the same as in (C).

N47H and G81A mutations in the PD and V265F in the HD and showed the highest recovery rates and possessed the largest freely diffusing pool of all the PAX3-GFP proteins (Fig. 4-8C & D). The remaining mutants (*Sp*^{d}, F45L, S84F, Y90H, and R271G), which we have called class II, displayed recovery rates and free pools intermediate between wild type PAX3 and class I mutants (Fig. 4-8C & D), with *Sp*^{d} being the fastest and Y90H the slowest amongst this cohort. As noted above, these differences in mobility also correlate with localization, where the class I mutants displayed a diffuse distribution that included a cytoplasmic fraction, while the class II mutants retained clear evidence of subnuclear compartmentalization. The specificity of these findings is underscored by the properties of the only two WS missense mutations outside of either DNA-binding domain, A196T and Q391H; consistent with the idea they occur at intron-exon junctions and are thought to contribute to PAX3 dysfunction by aberrant splicing, neither mutant had an appreciable affect on PAX3 mobility (Fig. 4-8C). Importantly, the differences in mobility among the two mutant classes and wild type PAX3-GFP are independent of protein expression levels. Specifically, none of the PAX3 variants were expressed at abnormally high levels, indicating that the increased mobility and unbound fraction of the fastest-recovering PAX3 variants is not due to the saturation of binding sites or other nuclear constituents that constrain PAX3 movement. Together, these observations indicate that disease-causing alleles of PAX3 exert significant effects on mobility and provide the first measure of PAX3 activity that is consistently altered in this collection of mutations.

The segregation of PAX3 mutants into two discrete categories based on localization and dynamics was unanticipated. To address whether these reflect

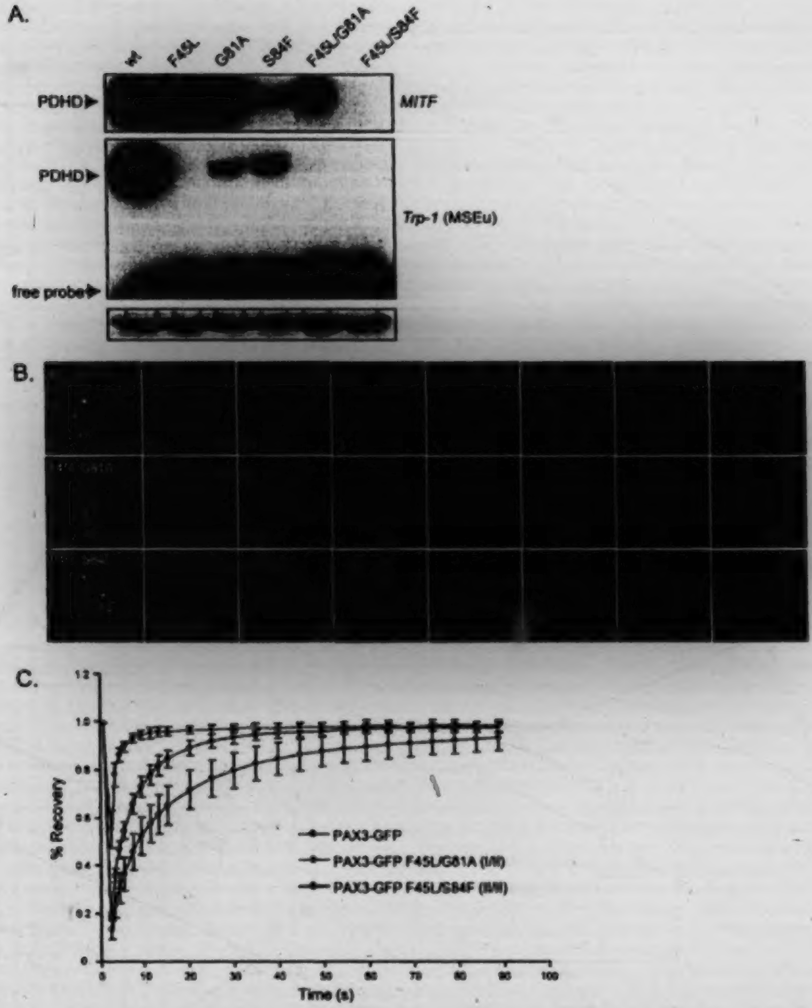


Figure 4-9. Mutations causing compartmentalization defects exert a dominant effect on PAX3 behavior in the nucleus. (A) Effects of class I (G81A), class II (F45L, S84F), class I/II (F45L/G81A), or class II/II (F45L/S84F) mutations on DNA binding properties of a recombinant protein encompassing the PAX3 paired domain and homeodomain (PDHD) were analyzed in mobility shift assays. Affinity of each mutant for the PAX3-binding elements of the *MITF* (*top panel*) and *Trp-1* (*middle panel*) promoter sequences was compared to the wild type PDHD protein (lane 1). Free probe is shown for the bottom panel only; all PDHD proteins were equalized to the same concentration, as indicated in the Coomassie gel (*bottom panel*). (B) FRAP experiments showing recovery properties of wild type PAX3-GFP (*top row*) and mutant PAX3-GFP proteins containing the F45L and G81A (*middle row*) or F45L and S84F (*bottom row*) mutations. (C) FRAP recovery curves for wild type and double mutant PAX3-GFP in 10T1/2 cells. Plots show recovery versus time; wild type PAX3-GFP plot represents the average of thirty individual FRAP experiments; mutant plots represent the average of ten FRAP experiments.

completely separate effects, and to further refine their mechanistic basis, representative mutants from each class were combined. This involved the creation of F45L/G81A (class II/class I) and F45L/S84F (class II/class II) PAX3-GFP fusions. In each case, DNA binding of complementary PDHD proteins was assessed and compared to wild type PAX3 and the singly mutated counterparts (F45L, G81A, and S84F). As expected, the presence of a class II mutation was always associated with a loss or reduction in DNA binding compared to wild type (Fig. 4-9A), and the class II/class II combination did not display defects in localization or mobility over and above each individual mutation (Fig. 4-9B & C). The class II/class I double mutant, however, fully recapitulated the mobility and localization of the class I mutant, but now with reduced (*MITF*) or absent (*Trp-1*) DNA binding. From a mechanistic standpoint, this indicates that the retention of DNA binding by class I proteins is not required for their increase in mobility or diffuse distribution and that class I mutations exert a dominant effect on these parameters of PAX3 activity. Moreover, it clearly establishes that class I and II mutations lead to PAX3 dysfunction via distinct means, such that altered nuclear dynamics appears to be the primary defect of the N47H, G81A, and V265F proteins.

4.9 Discussion

In this chapter, we have examined the subnuclear distribution and dynamics of the developmentally important transcription factor PAX3 and numerous disease-causing mutants. We find that PAX3 overlaps with a small number of discrete foci marked by transcriptional competence/activity, indicative of its localization within the interchromatin compartment and spatial proximity to the euchromatic compartment.

Using GFP fusion proteins, we have shown that PAX3 contains separable determinants for subnuclear compartmentalization and that DNA binding activity does not direct localization within the nucleus. Characterization of mutant PAX3-GFP proteins demonstrates that both subnuclear localization and dynamics of PAX3 are perturbed by mutations in the DNA-binding domains, likely influencing PAX3 function *in vivo*. In addition, we can divide our cohort of PAX3 disease alleles into two classes based on their localization and dynamics in the nucleus and demonstrate a relationship between intranuclear mobility and DNA binding that relates to molecular interactions which control PAX3 function *in vivo*.

Immunofluorescence of endogenous and exogenous PAX3 revealed a reticular distribution that was excluded from regions of condensed chromatin (Fig. 4-1) but juxtaposed to euchromatin (Fig. 4-2), consistent with localization to the interchromatin compartment. This was most apparent with the global proximity of PAX3 to epigenetic marks of transcriptionally active chromatin, notably H3K4me3 and H3K36me3, which are associated with promoter (Santos-Rosa et al. 2002; Ng et al. 2003) and transcribed regions (Krogan et al. 2003; Li et al. 2003; Xiao et al. 2003), respectively. In this regard, the distribution of PAX3 resembles that described for proteins that localize to the interchromatin space, including the general transcription factor TFIIH (Grande et al. 1997; Verschure et al. 2002) and a number of sequence-specific transcription factors (Spector 2003; see Section 1.3.2). Moreover, PAX3 mobility (t_{50} of ~7.9s) is similar to many other DNA binding and chromatin-associated proteins with half maximal recovery times on the order of several seconds (Phair et al. 2004). Lastly, the fact that only a small subset of PAX3 co-localizes with H3K4me3 or sites of transcription delineated by 5'-

FUrd incorporation also correlates with studies that showed only a fraction of the total nuclear transcription factor population coincided with foci formed by RNAPII (van Steensel et al. 1995; Elefanty et al. 1996; van Steensel et al. 1996; Grande et al. 1997). As a result, our findings with PAX3 are consistent with models where transcription factors accumulate predominantly in subnuclear domains within the interchromatin compartment, which may modulate their local concentration or serve as assembly sites for regulatory complexes and allow dynamic exchange with target sites (Hendzel et al. 2001; Hager et al. 2002; van Holde and Zlatanova 2006).

Disease mutations have been shown to disrupt or alter the *in vivo* action and *in vitro* biochemical properties of numerous transcription factors (Nazareth et al. 1999; Enwright et al. 2003; Black et al. 2004; Kino et al. 2004; Rajaram and Kerppola 2004; Sharp et al. 2004; see Sections 1.3.3 and 1.3.4). In the case of *PAX3*, disease-causing alterations range from complete deletions of the locus to single nucleotide substitutions. These lesions can lead to an absence of protein product, internal deletions or truncations, or missense mutations, which together establish that the WS phenotype arises from haploinsufficiency. With only two exceptions, A196T and Q391H, all known missense mutations occur in one of the two PAX3 DNA-binding domains (Read and Newton 1997; Fig. 1-4), and loss-of-function analyses have therefore focused on DNA binding and reporter gene assays. These studies have revealed a range of effects on PAX3 DNA binding and reporter gene activity and demonstrate that these parameters do not always correlate (Chalepakis et al. 1994a; Underhill et al. 1995; Epstein et al. 1996; Underhill and Gros 1997; Fortin et al. 1998; Watanabe et al. 1998; Corry and Underhill 2005a; see Chapter 3). Given this disparity, it is significant that disease alleles were found to have more consistent effects on PAX3 mobility and that these correlate with localization and DNA binding (discussed below). Specifically, all mutations present within the PD or HD were characterized by increased mobility, which is further underscored by the fact that A196T and Q391H resembled wild type PAX3 with respect to localization and mobility. This latter finding was not unexpected, as both mutations occur at intron-exon junctions and have been suggested to disrupt PAX3 function through aberrant splicing. Based on the fact that increased mobility is a shared feature amongst a randomly selected cohort of seven WS mutations (and one *Sp* mutation), we suggest this may be a general characteristic of disease-causing missense alleles.

The increase in the mobility of PAX3 mutants reflects a loss of molecular interactions that constrain movement of the wild type protein, the severity of which led to stratification of mutants into two groups. Accordingly, the mutants exhibiting the greatest mobility (class I: N47H, G81A and V265F) can be considered to manifest the most severe defect in these interactions and is consistent with their diffuse localization and lack of compartmentalization. Likewise, the intermediate effect of class II mutants (*Sp*^{d}, F45L, S84F, Y90H, and R271G) on PAX3 movement suggests they retain sufficient interactions to confer a compartmentalized appearance. In this regard, PAX3 is similar to other transcription factors, where altered nuclear distribution (Day et al. 2003; Enwright et al. 2003; Black et al. 2004; Rajaram and Kerppola 2004; Kawate et al. 2005) and increased mobility (Schaaf and Cidlowski 2003; Farla et al. 2004; Karpova et al. 2004; Schaaf et al. 2006) have been seen upon mutation of the DNA-binding domain, and are thought to derive from defects in protein-DNA interactions. For instance, analyses of wild type and mutant forms of the NF-κB subunit p65 found that high mobility correlated with low DNA binding affinity and a more random distribution in the nucleus (Schaaf et al. 2006), and an analogous relationship was also demonstrated for the GR (Schaaf et al. 2005). Similarly, disease-causing mutations in the PIT1 transcription factor lead to reductions in DNA binding activity and increased nuclear mobility (Sharp et al. 2004). A key difference in the behavior of PAX3 mutants, however, is that their DNA binding properties are inversely correlated with their subnuclear localization and mobility, since intranuclear behavior is most severely affected by mutations (class I) that permit interaction with target sites at levels similar to wild type PAX3. Although this suggests that DNA binding might be required for the rapid mobility and diffuse distribution of class I mutants, these attributes are retained even with the loss or reduction of DNA binding that occurs by combining with a class II mutation (Fig. 4-9). Together, these data indicate altered localization and mobility are important aspects of PAX3 dysfunction in WS, and represent the principal defect in class I mutants where they supersede DNA binding.

Class I and II mutants can be separated by a single amino acid in the PAX3 primary structure, as seen for F45L and N47H. Moreover, when these mutants are examined in the context of the DNA-bound PD or HD crystal structures, or models where the domains are bound to the composite *PH0* motif (Jun and Desplan 1996), there is no obvious correlation between the position of mutations and their effects on localization and mobility (summarized in Fig. 4-10A). For instance, mutations in separate DNA-binding domains could belong to the same mutant class, while those that were spatially close were just as likely to segregate to different classes. This can be illustrated by the proximity of F45L (class II) and N47H (class I), which both project side chains towards

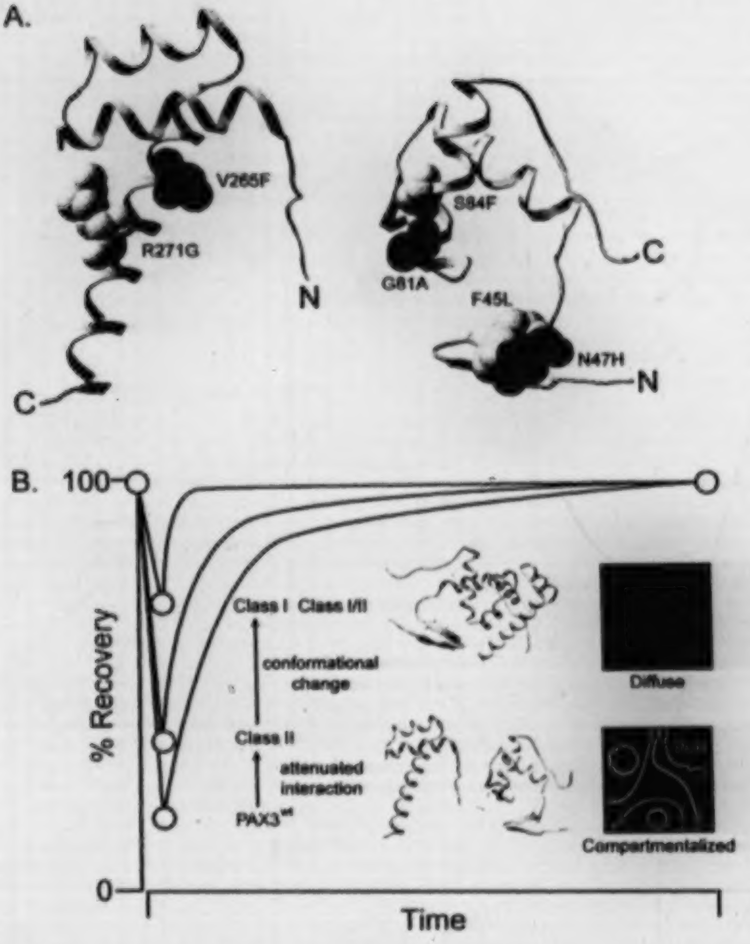

A.
V265F
R271G
N
C
S84F
G81A
F45L
N47H
C
N
B.
100
0
% Recovery
Time
Class I
Class I/II
conformational change
Class II
attenuated interaction
PAX3wt
Diffuse
Compartmentalized

Figure 4-10. Disease mutations exert differential effects on determinants that constrain PAX3 activity in the nucleus. (A) Ribbon structures of the PAX3 homeodomain (*left*) and amino-terminal paired subdomain (*right*) showing the spatial proximity of selected class I (*dark grey*) and class II (*light grey*) mutations. The model depicts the DNA binding surface of the PD and HD oriented as they would appear on a composite *PH0* motif. (B) Summary of the effects of class I and II mutations on PAX3 behavior *in vivo*. Wild type PAX3 and proteins from the two mutant classes each display unique recovery profiles in FRAP experiments; the increase in mobility among class II mutants reflects impairments in molecular interactions that constrain local movement of PAX3, while conformational alterations (indicated by the change in structure, *middle*) appear to underlie the further increase in mobility of class I mutants. Schematic diagrams on the right summarize the compartmentalized pattern of class II mutants (*bottom*) and diffuse, uncompartmentalized phenotype of class I mutants (*top*). Green, PAX3; red, Hoechst; black, pericentromeric heterochromatin foci (PCH).

the minor groove in the paired domain crystal structure, while G81A (class I) and S84F (class II) lie on the same surface of helix 3 of the amino-terminal PD subdomain, facing the major groove (Fig. 4-10A). As a result of this spatial relationship, it is unlikely that each mutational class affects a distinct interaction surface. Rather, these data suggest that class II proteins retain determinants that constrain their localization, but attenuation of the underlying interaction(s) that support compartmentalization leads to increased mobility (Fig. 4-10B). In contrast, class I mutations would cause a loss of key interaction surfaces altogether and can act in a dominant manner to induce rapid mobility and diffuse localization when combined with a class II mutation (Fig. 4-10B). The most likely explanation is that class I mutations alter PAX3 conformation and is in keeping with previous studies that established important roles for protein chaperones in regulating transcription factor mobility, as well as their transport to and from the nucleus (reviewed in Hager et al. 2006). It is therefore noteworthy that class I mutants exhibit a cytoplasmic fraction, reflecting a defect in nuclear import or retention, and highlighting further differences with class II mutants. Regardless, the increase in mobility of PAX3 mutants represents a thermodynamic defect that will impede their ability to form functional complexes during transcription, an idea that is supported by the dynamic interdependence of the GR and HMGB1 proteins (Agresti et al. 2005). Together, these data indicate that mutations in close proximity in both the primary and tertiary PAX3 structure can have remarkably diverse effects. Finally, both altered mobility and conformation are likely to introduce a stochastic element to PAX3 function that may contribute to phenotypic diversity in Waardenburg syndrome.

CHAPTER 5. DETERMINANTS OF PAIRED-TYPE HOMEODOMAIN BEHAVIOR

Structural modeling of HD surfaces in Fig. 5-1 was performed by Dr. Alan Underhill; bioinformatic analyses of HD-binding motifs in Fig. 5-4 was performed by Dr. Alan Underhill; FRAP data in Fig. 5-5 was collected and compiled by Kristal Missiaen.

5.1 Background

Group III and IV PAX proteins are distinct from the other family members due to the presence of the HD, which can cooperatively bind to DNA with the PD or function as an independent DNA-binding module when isolated from the full length protein. The presence of the HD in *Drosophila* Prd and its conservation in four of the nine PAX family proteins (Fig. 1-2) suggests that its function is required for some aspect(s) of these proteins' behavior. In support, functional interdependence between the PD and HD has been demonstrated (Underhill et al. 1995; Underhill and Gros 1997; Apuzzo and Gros 2002; Apuzzo et al. 2004) and human disease-causing mutations in the PAX3 (Fortin et al. 1997; Corry and Underhill 2005a) and PAX6 (Singh et al. 2000; Mishra et al. 2002) HDs are known to interfere with both PD and HD DNA binding activity. However, mutational studies have suggested that the HDs of PAX6 and eyeless (the *Drosophila* ortholog of PAX6) exert minimal influence on the protein's overall function (Punzo et al. 2001; Haubst et al. 2004). Furthermore, homeoproteins involved in *Drosophila* development in which the HD has been deleted have nevertheless been shown to rescue mutants and properly regulate embryo patterning (Fitzpatrick et al. 1992; Copeland et al. 1996; Punzo et al. 2001). Thus, despite its presence in many transcription factors, the contribution of the HD to these proteins' functions remains poorly understood.

There has also been a long-standing question of how homeobox transcription factors achieve functional specificity using a simple target motif (5'-TAAT) that occurs regularly in the genome (Hayashi and Scott 1990). Based on studies of *Drosophila* homeoprotein behavior, several models have been put forward to account for the discrepancy between the promiscuity of target recognition *in vitro* and precise control of target gene expression *in vivo* (reviewed in Biggin and McGinnis 1997). It is widely accepted that HD residue 50 is the major determinant of target specificity (Hanes and Brent 1989; Treisman et al. 1989); however, studies using chimeric homeoproteins have demonstrated the importance of regions both within and external to the HD for target discrimination and affinity (Hoey et al. 1988; Kuziora and McGinnis 1989; Gibson et al. 1990; Damante and Di Lauro 1991; Dessain et al. 1992; Ekker et al. 1992; Furukubo-Tokunaga et al. 1992; Lin and McGinnis 1992; Chan and Mann 1993; Furukubo-Tokunaga et al. 1993; Frazee et al. 2002).

Our prior analyses of PAX3 subnuclear compartmentalization suggested that the only contiguous segment that conferred discrete localization to a particular compartment (i.e., chromatin) was the HD. Here, we have carried out a comparative analysis involving the PAX3 HD and two other Prd-type HDs to determine the underlying basis of this activity and to assess whether it is specific to the PAX3 HD. In comparing the PAX3 HD and related HDs from the PRRX1 and PITX2 transcription factors, we show that, despite exhibiting extensive overlap in their target specificities, all three adopt unique subnuclear localization patterns. We also undertake a detailed examination of the relationship between HD DNA binding and localization, and demonstrate that mutation of residue 50 not only alters target specificity *in vitro* but also affects subnuclear localization, although DNA binding preference appears to have little influence over HD localization in the nucleus. Together, our observations suggest that multiple determinants are responsible for regulating the behavior of Prd-type HDs and that variations in the overall structure of the HD may influence both subnuclear localization and DNA binding.

5.2 Comparison of the PAX3, PRRX1, and PITX2 homeodomains

The Prd subclass of proteins feature HDs that share homology with those of the *Drosophila* Prd, Gsb-proximal and Gsb-distal transcription factors (Bopp et al. 1986; see Section 1.1.3). The Prd subclass can be further divided into three groups, depending on the presence of a serine, glutamine, or lysine at position 50 of the HD. Most HDs interact with sequences containing a 5'-TAAT core motif (Kalionis and O'Farrell 1993). Residue 50, located in the recognition helix, specifies the two bases on the 3' side of this motif (Hanes and Brent 1989; Treisman et al. 1989), allowing HDs to recognize similar but distinct targets. Thus, HDs from the three Prd-type classes are expected to interact with different target sequences, and it has been shown that derivatives of the Prd HD exhibit preferences for a range of sequences, depending on the residue at position 50 (Wilson et al. 1993). We previously analyzed several Prd-type HDs for their ability to cooperatively bind to the *MITF* promoter sequence with the PAX3 PD (see Fig. 3-10). Two of these HDs, from the PRRX1 and PITX2 proteins, are members of the Gln50 and Lys50 Prd-type HD subclass, respectively. The PRRX1 and PITX2 HDs share 67% and 62% sequence identity, respectively, with the PAX3 HD (Fig. 3-10A), but are characterized by extensive homology in the recognition helix (Fig. 5-1A). Although the PAX3 (10.74), PRRX1 (11.65), and PITX2 (11.54) HDs all possess basic isoelectric points, each protein varies in the ionic charge projected from its solvent-exposed surface (Fig. 5-1B), which may affect interactions with proteins in the surrounding environment.

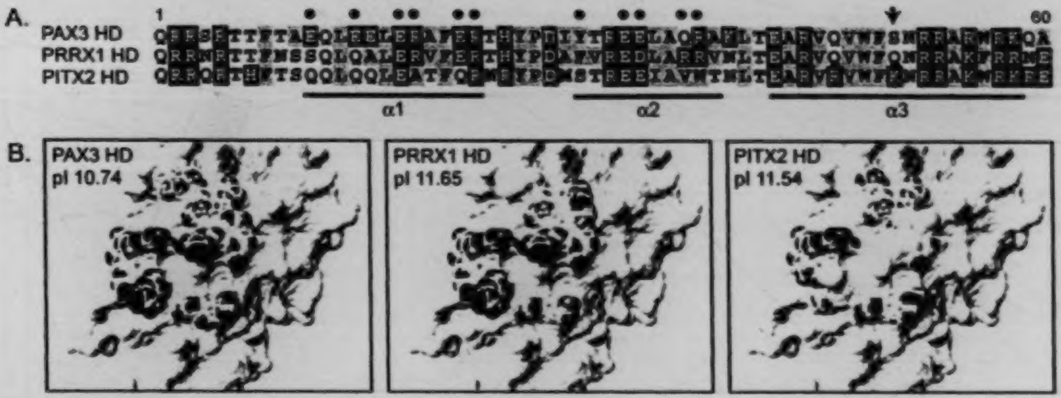

A.
1
60
PAX3 HD
PRRX1 HD
PITX2 HD
α1
α2
α3
B.
PAX3 HD
pI 10.74
PRRX1 HD
pI 11.65
PITX2 HD
pI 11.54

Figure 5-1. Comparison of the PAX3, PRRX1, and PITX2 homeodomains. (A) Alignment of the PAX3, PRRX1, and PITX2 HD amino acid sequences. Residues are color-coded as follows: hydrophobic (grey), polar uncharged (yellow), acidic (red), basic (blue). Residue 50 is indicated by an arrow; residues that form the solvent-exposed surface are indicated by filled circles; sequences that form alpha-helices are underlined. (B) Structural modeling of the solvent-exposed surfaces of the PAX3, PRRX1, and PITX2 HDs. Hydrophobic (grey), acidic (red), and basic (blue) residues that project from helices 1 and 2 of the HDs are shown and isoelectric points (pI) are indicated.

5.3 Subnuclear localization of the PAX3, PRRX1, and PITX2 homeodomains

In the previous chapter, we demonstrated that the GFP-tagged PAX3 HD preferentially localizes to condensed chromatin (Fig. 4-4C). This was a surprising result, considering that we never observed interaction of endogenous PAX3 with heterochromatic regions at appreciable levels. Thus, we wished to examine the PAX3 HD in greater detail to determine if its localization would reveal information about its activity that is relevant to PAX3 function, and also to investigate whether related HDs display similar *in vivo* behavior. We tagged the PRRX1 and PITX2 HDs with GFP and, with the PAX3 HD-GFP protein, observed their localization in 10T1/2 cells. PRRX1 HD-GFP did not co-localize with pericentromeric heterochromatin to the same extent as PAX3 HD-GFP (Fig. 5-2A, *middle row*), although some overlap with areas of Hoechst stain was observed, especially on the nucleolar periphery. The pattern of PITX2 HD-GFP was unique among the three proteins, most notably due to its prominent nucleolar localization (Fig. 5-2A, *bottom row*). In addition, the majority of PITX2 HD-GFP was excluded from condensed chromatin, with minimal overlap in pericentromeric regions. These results indicate that the association between PAX3 HD and condensed chromatin is specific and also make the important observation that closely related HD proteins are characterized by distinct subnuclear localization patterns. Interestingly, the intranuclear distributions of the full length PAX3, PRRX1, and PITX2 proteins are almost identical, as each protein formed a reticulated pattern and was excluded from Hoechst-rich areas (Fig. 5-2B). These observations confirm that regions flanking the HD exert influence on behavior of the HD *in vivo* and that, in these proteins at least, any role the HD plays in subnuclear localization is suppressed by other determinants.

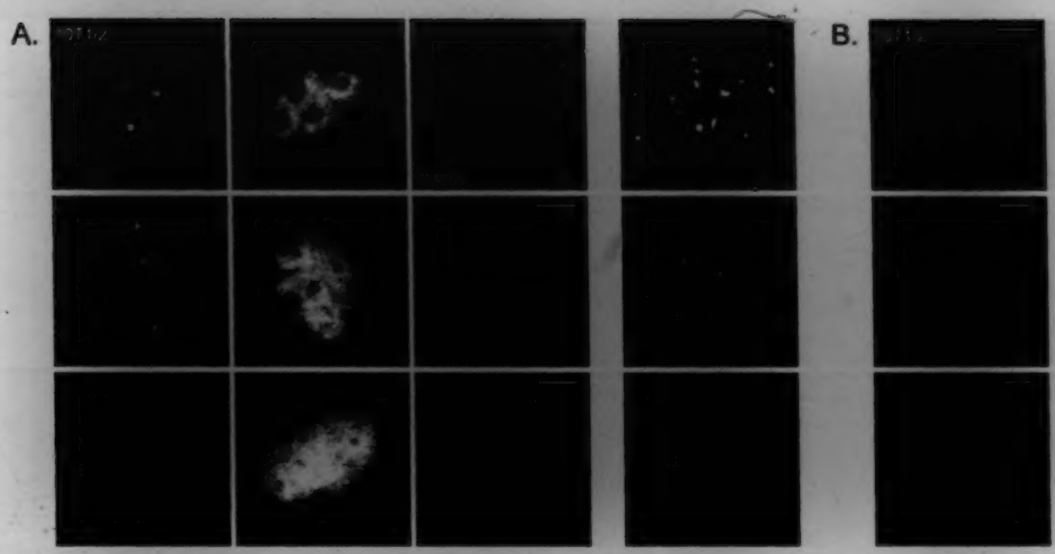


Figure 5-2. Comparison of the subnuclear localization of the PAX3, PRRX1, and PITX2 homeodomains. (A) Subnuclear localization patterns of the PAX3 (*top row*), PRRX1 (*middle row*), and PITX2 (*bottom row*) HDs expressed as GFP fusions in 10T1/2 cells. Signal intensity thresholds (*far right column*) detail regions of overlap (yellow) between Hoechst (red) and GFP (green). Localization patterns of full length GFP-tagged PAX3 and PRRX1 and HA-tagged PITX2 in 10T1/2 cells (red = Hoechst) are shown for comparison (B). Grayscale panels are unaltered; merged panels have been deconvolved. Bar, 10 μm.

5.4 Analysis of Paired-type homeodomain target preference

A key determinant of HD target binding specificity is the identity of residue 50 (Hanes and Brent 1989; Treisman et al. 1989). Each of the PAX3, PRRX1, and PITX2 HDs binds preferentially to sequences containing a 5'-TAAT core motif (Grueneberg et al. 1992; Underhill and Gros 1997; Amendt et al. 1998); however, differences at position 50 among these HDs specify unique combinations of bases immediately 3' to the TAAT motif (Fig. 5-3A). We therefore wished to assess if the distinct target specificity of the three HDs accounted for their unique subnuclear localization patterns. We began by examining the DNA binding behavior of the three HDs through a series of DNA binding assays with oligonucleotides containing the preferred binding targets of the three HDs. As Fig. 5-3B shows, the PAX3, PRRX1, and PITX2 HDs display a great deal of overlap in sequence recognition. Remarkably, the PAX3 HD interacted with all three target sequences (*lane 1*), as did the PRRX1 HD (*lane 4*), although with minimal affinity to the *bicoid* sequence. The PITX2 HD possessed the most restricted specificity, interacting only with its preferred sequence (*lane 7*). We also assessed binding of the three wild type HDs to the *P2* oligonucleotide, which mediates cooperative dimerization of Prd-type HDs (Wilson et al. 1993). The PAX3 HD dimerized efficiently on *P2*; however, the PITX2 HD showed only weak monomeric binding and, while the PRRX1 HD displayed high affinity binding as a monomer, it was limited in its ability to dimerize on the sequence (Fig. 5-3C). These results are consistent with the findings of Wilson et al. (1993), which showed that the Ser50 Prd HD cooperatively dimerizes on the *P2* sequence, but variants in which a Gln or Lys were substituted at position 50 favor TAAT half sites separated by 3 bp and prefer different combinations of bases that space the two half sites. These

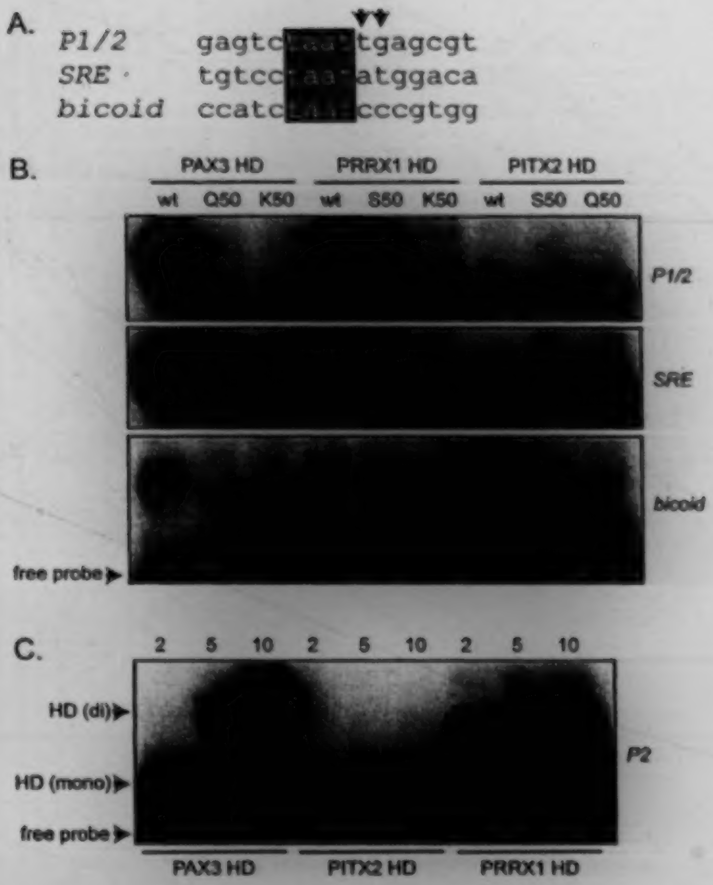

A.
P1/2 gagtc taat tgagcgt
SRE tgtcc taat atggaca
bicoid ccatc taat cccgtgg
B.
PAX3 HD
PRRX1 HD
PITX2 HD
wt Q50 K50 wt S50 K50 wt S50 Q50
P1/2
SRE
bicoid
free probe
C.
2 5 10 2 5 10 2 5 10
HD (di)
HD (mono)
free probe
P2
PAX3 HD
PITX2 HD
PRRX1 HD

Figure 5-3. Homeodomains from the PAX3, PRRX1, and PITX2 proteins display overlapping target specificity. (A) Alignment of the preferred binding sequences of the PAX3 (*P1/2*), PRRX1 (*SRE*), and PITX2 (*bicoid*) HDs. The core TAAT motif is shaded in black; residues specified by position 50 are indicated with arrows. (B) Mobility shift assays used to address interaction of the wild type PAX3, PRRX1, and PITX2 HDs, or variants in which position 50 was substituted, with the *P1/2* (top), *SRE* (middle), and *bicoid* (bottom) oligonucleotides. Position of the free probe is shown for the bottom panel only. (C) Cooperative binding of the PAX3, PRRX1, and PITX2 HDs to the *P2* oligonucleotide. Amounts (µL) of each protein used are indicated above each lane. Positions of the HD-DNA (*mono* = monomer; *di* = dimer) complexes and free probe are indicated.

results confirm that residue 50 plays an important, but not absolute, role in HD target specificity and support data showing that it influences Prd-type HD cooperativity.

To gain more information regarding the role of residue 50 in target specificity, we exchanged the amino acid at position 50 among the PAX3, PRRX1, and PITX2 HDs and performed DNA binding assays. As expected, mutation of residue 50 conferred altered target specificity to all the HDs, but also modified other aspects of PAX3 and PRRX1 HD DNA binding behavior (Fig. 5-3B). Introduction of a Gln at position 50 severely reduced DNA binding by the PAX3 HD (*lane 2*), while a Ser50→Lys mutation prohibited PAX3 HD binding to the *P1/2* and *SRE* sequences but appeared to slightly improve affinity for the *bicoid* site (*lane 3*). In addition, the PAX3 Lys50 HD-*bicoid* complex showed increased mobility compared to that formed by the wild type PAX3 HD, suggestive of an altered conformation. Target preference of the PRRX1 Ser50 HD did not change compared to wild type PRRX1 HD (*lane 5*); however, two HD-DNA complexes were detected in the presence of the *P1/2* and *SRE* oligonucleotides, again suggesting altered conformation of the HD as a result of substituting residue 50. Mutating Gln50 to Lys did not alter PRRX1 HD interaction with the *P1/2* or *SRE* sequences, but supported binding to the *bicoid* sequence, which was not possible with the wild type or Ser50 variants. Finally, mutation of Lys50 of the PITX2 HD to Ser or Gln abrogated interaction with the *bicoid* sequence, but switched its target preference for the PAX3 HD- and PRRX1 HD-preferred sites. These results clearly demonstrate the importance of residue 50 in HD-target interaction, but also reveal a degree of flexibility in DNA binding preference by related HDs. More importantly, our observations show that altering HD target specificity is not exclusively dependent on residue 50, suggesting that other amino acids in the HD influence DNA binding and target preference.

5.5 Mutation of residue 50 alters HD intranuclear compartmentalization

Having established that substitution of amino acid 50 can alter HD target specificity, we next assessed how mutation at this position affects subnuclear localization. To test this, we created GFP fusions of the PAX3, PRRX1, and PITX2 HDs containing a Ser, Gln, or Lys at position 50. As shown previously, wild type PAX3 HD-GFP displayed considerable overlap with Hoechst-stained chromatin, particularly in perinucleolar and pericentromeric regions (Fig. 5-4A, *top row*). In contrast, PAX3 HD-GFP Gln50 (Fig. 5-4A, *middle row*) and Lys50 (Fig. 5-4A, *bottom row*) displayed reduced localization to heterochromatic areas. The Lys50 mutation exerted the most severe effect, almost completely relocalizing the PAX3 HD to non-heterochromatic regions, but failed to cause nucleolar localization, as we had observed with the PITX2 HD-GFP protein (Fig. 5-2A). PRRX1 HD-GFP, as before, was distributed throughout the nucleus and showed minimal overlap with condensed chromatin (Fig. 5-4B, *top row*). PRRX1 HD-GFP Ser50 showed the same diffuse pattern as wild type, but displayed increased co-localization with Hoechst-stained regions (Fig. 5-4B, *middle row*). Substituting Gln50 for Lys also led to a diffuse distribution, but caused robust nucleolar localization and appeared to completely distribute the PRRX1 HD away from Hoechst-rich areas (Fig. 5-4B, *bottom row*). The PITX2 HD-GFP protein, as before, was distributed throughout the nucleus, showing strong nucleolar localization and moderate overlap with perinuclear chromatin (Fig. 5-4C, *top row*). The Ser50 variant showed

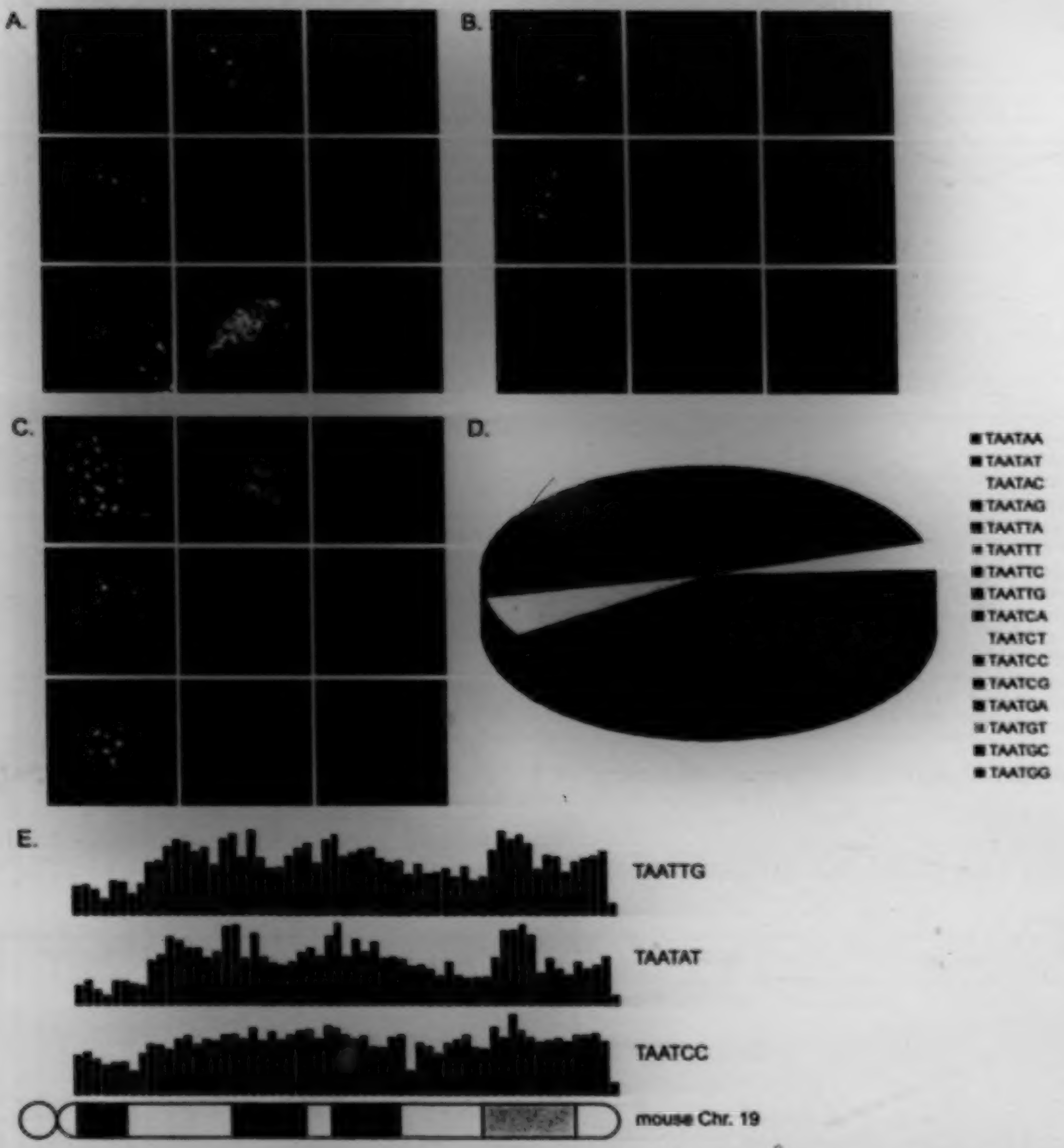


Figure 5-4. Mutation of residue 50 affects homeodomain subnuclear localization. Effects of exchanging HD residue 50 among the PAX3 (A), PRRX1 (B), and PITX2 (C) HDs was examined by observing subnuclear distribution of wild type (*top rows in A, B, & C*) and mutant GFP-tagged HDs. Grayscale panels are unaltered; merged panels have been deconvolved. Bar, 10 µm. (D) Distribution of all possible HD-binding motifs (5'-TAATNN) throughout the mouse genome. Frequencies were determined without repeat-masking. (E) Distribution of preferred binding motifs for the PAX3 (TAATTG, *top*), PRRX1 (TAATAT, *middle*), and PITX2 (TAATCC, *bottom*) HDs along mouse chromosome 19 (*shown schematically at bottom*). Each bar represents one megabase (Mb) of DNA; density of TAATNN sites within each Mb segment correlates with bar height.

reduced nucleolar localization, instead occupying primarily the inner nucleolar periphery. Significantly, the PITX2 HD-GFP Ser50 co-localized with condensed chromatin in both pericentromeric and perinuclear regions (Fig. 5-4C, *middle row*), much like wild type PAX3 HD-GFP. Finally, the PITX2 HD-GFP Gln50 protein was still found in the nucleolus but displayed a more diffuse pattern than the wild type version and exhibited limited overlap with pericentromeric regions, as observed for PRRX1 HD-GFP (Fig. 5-4C, *bottom row*).

We also analyzed the frequencies of all sixteen possible 5'-TAATNN motifs in the mouse genome to ensure that the subnuclear localization patterns of the PAX3, PRRX1, and PITX2 HDs were not a result of interaction with sites enriched in certain regions of the genome. Our results demonstrated that no individual TAATNN motif is enriched or over-represented in the mouse genome (Fig. 5-4D), reducing the likelihood that target preference dictates subnuclear localization of HDs. We also analyzed the linear distribution of the preferred target motifs of the PAX3, PRRX1, and PITX2 HDs along a representative mouse chromosome (chromosome 19) (Fig. 5-4E). Although subtle differences exist between the distributions of the three sites, we did not observe any elevation in frequency of these sites in any region of the chromosome, suggesting that the subnuclear localization of HDs is not due to local enrichments of preferred target motifs along chromosomes. These results suggest that HD residue 50 may exert other effects, over and above target preference, that influence compartmentalization in the nucleus.

5.6 Intranuclear mobility of wild type and residue 50-mutant Paired-type homeodomains

In Chapter 4, we demonstrated that PAX3 HD-GFP displays a rapid recovery following photobleaching and possesses a relatively large unbound protein pool (Fig. 4-5). Here, we used FRAP to compare the dynamics of the GFP-tagged PAX3, PRRX1, and PITX2 HDs in live cells and also investigated the effects of mutating residue 50 on the subnuclear dynamics of the PAX3 HD. Despite the obvious differences in their subnuclear localization patterns in live and fixed cells, we observed little variation in mobility among the three wild type HDs (Fig. 5-5A & B). PITX2 HD-GFP displayed the lowest effective recovery and freely diffusing pool, while the PRRX1 and PAX3 HDs exhibited higher recovery rates. These results suggest that related HDs possess determinants that confer similar intranuclear dynamics and mobility, even though they assume unique compartmentalization patterns.

Mutating Ser50 to Gln caused no apparent change in the mobility of the PAX3 HD-GFP protein (Fig. 5-5C, *left*); however, the Ser50→Lys version displayed a reduced free pool and slightly slower recovery compared to the wild type protein (Fig. 5-5C, *right*). It is possible that differences in the DNA-binding properties of the Lys50 HDs account for their reduction in mobility and free protein pool, as structures of various Lys50 HDs show that this residue makes two hydrogen bonds with guanines on the complementary strand of the 5’-TAATCC motif (Tucker-Kellogg et al. 1997; Chaney et al. 2005), while substitution of Gln50 for Lys in the engrailed (En) HD caused a substantial increase in DNA-binding affinity (Ades and Sauer 1994).

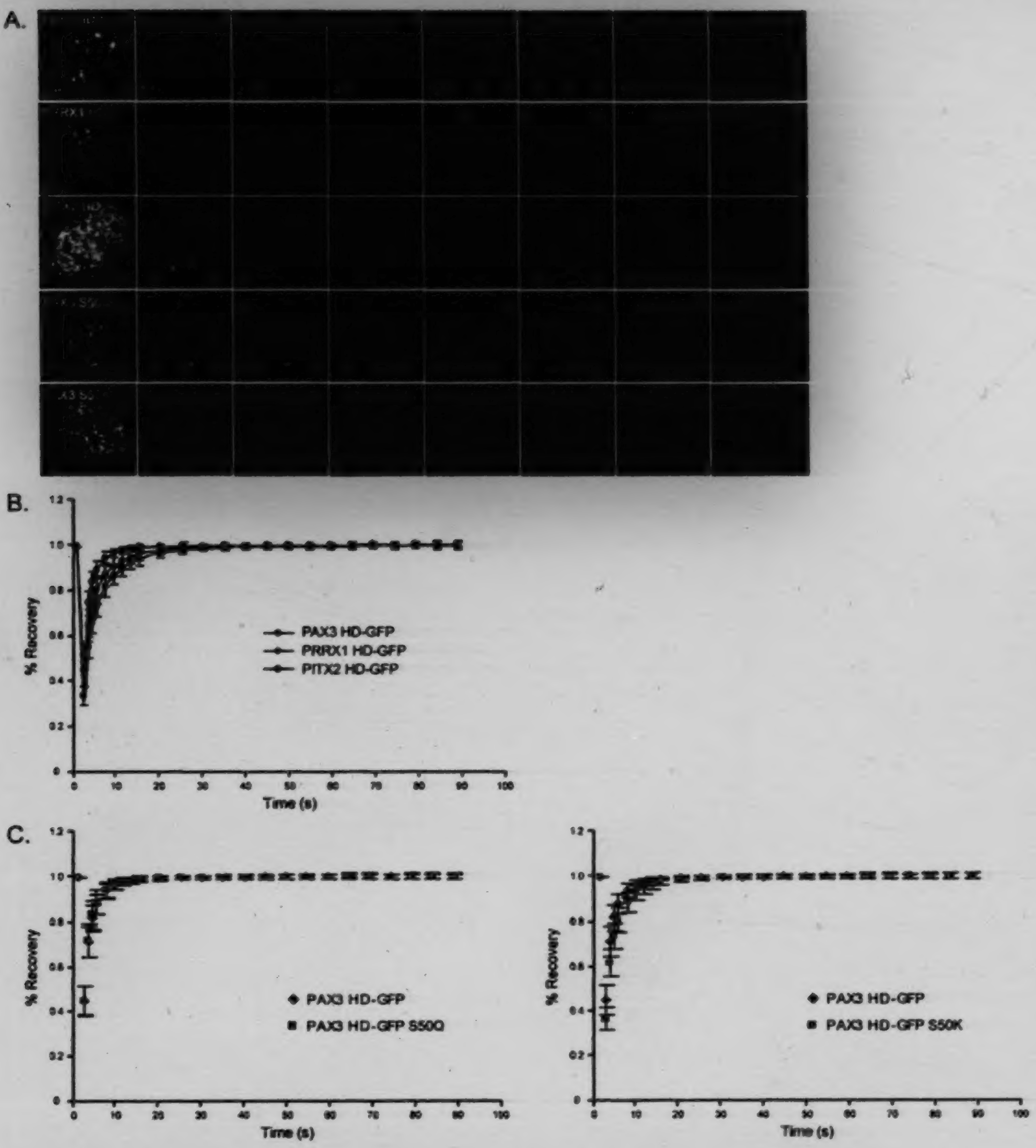


Figure 5-5. Intranuclear dynamics of wild type and residue 50-mutant homeodomains. (A) Localization of the wild type PAX3 HD-GFP, PRRX1 HD-GFP, and PITX2 HD-GFP proteins (*lanes 1-3*) and S50Q (*fourth row*) or S50K (*fifth row*) versions of PAX3 HD-GFP in live cells. (B) FRAP recovery curves for the wild type PAX3, PRRX1, and PITX2 HDs in 10T1/2 cells. (C) FRAP recovery plots for the PAX3 HD-GFP S50Q (*left*) and S50K (*right*) proteins (pink squares) compared to the wild type PAX3 HD-GFP recovery curve (blue diamonds). Plots show recovery versus time; each plot represents the average of thirty individual FRAP experiments. Error bars indicate standard deviation at each time point.

5.7 Structural compatibility of homeodomain residues influences DNA binding and subnuclear localization

Our data so far suggest that target preference plays a minor role in regulating the distribution of Prd-type HDs in the nucleus, indicating that amino acids other than residue 50 make important contributions to HD behavior *in vivo*. It is therefore possible that intrinsic properties, such as conformation or structure, modulate intranuclear compartmentalization of the HD, and that the same determinants influence HD target preference. Studies of altered-specificity HD mutants, such as the En Q50K protein (Tucker-Kellogg et al. 1997) and Asn51 mutants of OCT1 (Pomerantz and Sharp 1994), have shown that HDs can accommodate a limited number of amino acid identities at positions that dictate target preference and affinity, and that the overall conformation of the HD (and DNA structure) influences what residues can occupy these positions. Aside from residue 50, the PAX3, PRRX1, and PITX2 HDs share considerable sequence identity within their recognition helices, but show significant differences within helices 1 and 2, including residues on the solvent-exposed surface (Fig. 5-1). Depending on the identity of residue 50, these differences could account for the observed effects on HD DNA binding and subnuclear localization.

To address how residues outside the recognition helix affect HD DNA binding, we created chimeric proteins containing the recognition helix of the PAX3 HD fused to the first two helices of either the PRRX1 or PITX2 HDs (Fig. 5-6A, *top*) and compared their DNA binding behavior to that of the wild type PAX3 HD. The PRRX1-PAX3 HD did not bind to the *P1/2* sequence, contrasting the activity of the PAX3 and PRRX1 HDs (Fig. 5-6A, *lane 2*; *cf.* Fig. 5-2B). The PITX2-PAX3 HD bound to the *P1/2* site, albeit

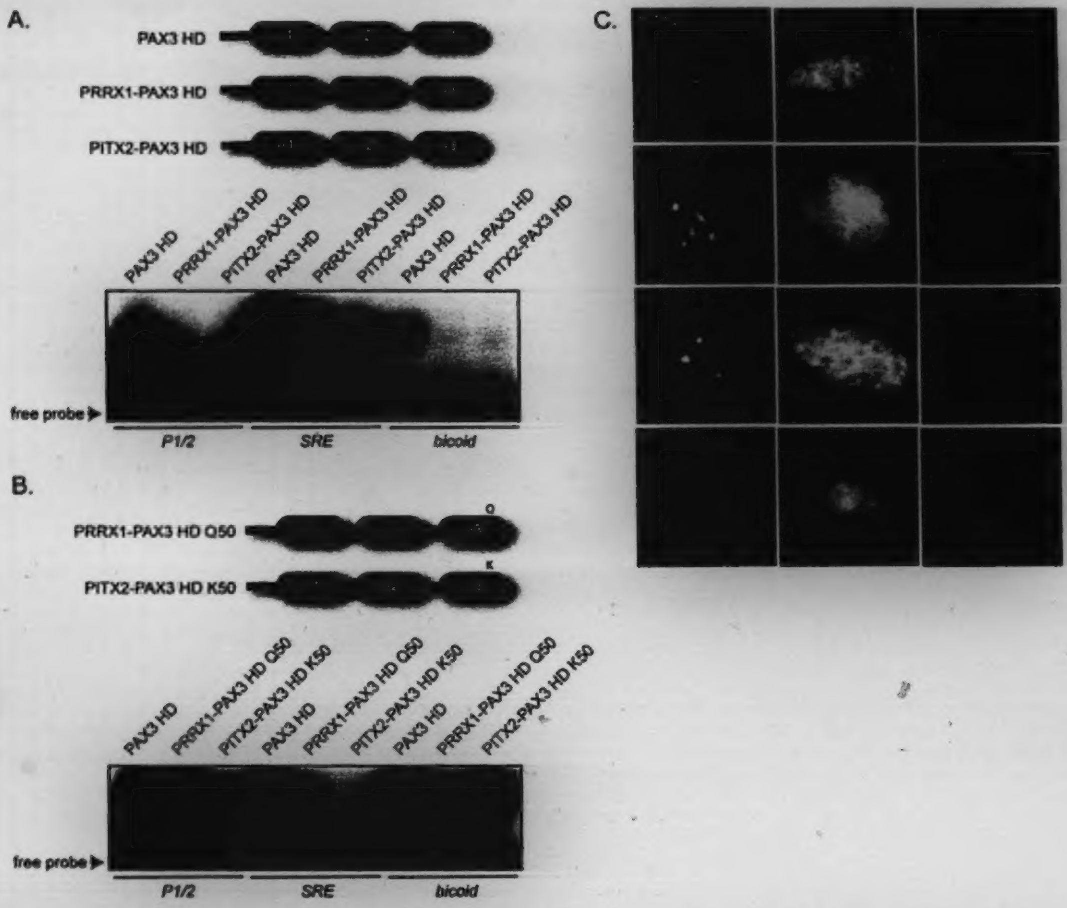

A.
PAX3 HD
PRRX1-PAX3 HD
PITX2-PAX3 HD
PAX3 HD
PRRX1-PAX3 HD
PITX2-PAX3 HD
PAX3 HD
PRRX1-PAX3 HD
PITX2-PAX3 HD
PAX3 HD
PRRX1-PAX3 HD
PITX2-PAX3 HD
free probe
P1/2
SRE
bicoid
B.
PRRX1-PAX3 HD Q50
Q
PITX2-PAX3 HD K50
K
PAX3 HD
PRRX1-PAX3 HD Q50
PITX2-PAX3 HD K50
PAX3 HD
PRRX1-PAX3 HD Q50
PITX2-PAX3 HD K50
PAX3 HD
PRRX1-PAX3 HD Q50
PITX2-PAX3 HD K50
free probe
P1/2
SRE
bicoid
C.

Figure 5-6. Contribution of regions outside the recognition helix to homeodomain behavior. (A) *Top*, schematic of chimeric HD constructs in which the recognition helix of the PRRX1 (green) or PITX2 (red) was replaced with that of PAX3 (blue). *Bottom*, binding assay used to assess interaction of the wild type and chimeric HDs with the *P1/2* (lanes 1-3), *SRE* (lanes 4-6), and *bicoid* (lanes 7-9) sequences. (B) *Top*, schematic of PRRX1-PAX3 and PITX2-PAX3 chimeric HD constructs (see above) in which Ser50 of the PAX3 recognition helix was exchanged for Gln or Lys, respectively. *Bottom*, binding assay used to examine affinity of the wild type PAX3 HD or residue 50-substituted chimeric HDs for the *P1/2* (lanes 1-3), *SRE* (lanes 4-6), and *bicoid* (lanes 7-9) sequences. (C) Intranuclear localization of the PRRX1-PAX3 S50 (*top row*) and Q50 (*second row*) and PITX2-PAX3 S50 (*third row*) and K50 (*bottom row*) HD-GFP proteins in 10T1/2 cells. Bar, 10 μm.

with reduced affinity compared to the wild type PAX3 HD (Fig. 5-6A, *lane 3*). Both the PRRX1-PAX3 (Fig. 5-6A, *lane 5*) and PITX2-PAX3 HDs (Fig. 5-6A, lane 6) interacted with *SRE*, although we observed different mobilities of the two complexes, suggestive of distinct binding conformations. In contrast to the wild type HD, neither chimeric HD bound to the *bicoid* sequence. The general trend in DNA binding among the chimeric HDs is a reduction in binding affinity, suggesting an incompatibility of residues in the PAX3 recognition helix with those in the PRRX1 or PITX2 HDs. This is surprising, considering the extensive sequence homology of the recognition helices from the three HDs, excluding position 50. Nevertheless, these results confirm that residue 50 is not the sole determinant of HD target specificity, and that other regions of the HD influence DNA binding behavior.

To clarify the influence of amino acids in helices 1 and 2 on the action of residue 50, we performed the same DNA binding assays with chimeric HDs in which Ser50 was substituted for Gln or Lys (Fig. 5-6B). Mutation of Ser50 to Gln fully recovered binding of the PRRX1-PAX3 HD to the *P1/2* sequence, but failed to improve affinity for the *SRE* or *bicoid* sites, whereas introduction of Lys50 into the PITX2-PAX3 chimera precluded binding to the *P1/2* and *SRE* oligonucleotides but facilitated high affinity binding to the *bicoid* sequence. Collectively, these results show that amino acids in the first two α-helices of the HD make important contributions to target recognition and DNA binding affinity. Given the obvious effect that substituting residue 50 had on DNA binding by the chimeric HDs, our data also indicate a requirement for compatibility between residues that specify target preference and those having no direct role in DNA contact or target

Table 7. Target preference of wild type, residue 50-mutant, and chimeric homeodomains.

HD variant	*P1/2*	*SRE*	*Bicoid*
PAX3 S50	++*	++	+
PAX3 Q50	-	-	-
PAX3 K50	+/-	+/-	+
PRRX1 Q50	++	++	+/-
PRRX1 S50	++[†]	++[†]	-
PRRX1 K50	+	+	+
PITX2 K50	-	-	+
PITX2 S50	+	+	-
PITX2 Q50	+	+	-
PRRX1-PAX3 S50	-	+	-
PITX2-PAX3 S50	+	+	-
PRRX1-PAX3 Q50	++	+/-	-
PITX2-PAX3 K50	-	-	++

*plus/minus signs indicate binding affinity; ++, high affinity; +, moderate affinity; +/-, minimal affinity; -, no binding observed.

[†] the PRRX1 S50 HD forms two distinct complexes on these sequences.

preference. The DNA binding affinity and specificity of all HDs examined is summarized in Table 7.

Next, we expressed equivalent PAX3 HD chimeras as GFP fusions in 10T1/2 cells to examine intrinsic determinants of HD localization in the nucleus. Both Ser50 chimeras were incompletely targeted to the nucleus (Fig. 5-6C; *top and third rows*), suggestive of a folding problem that interferes with nuclear import. The portion that did localize to the nucleus showed no discernable compartmentalization and did not overlap with Hoechst-stained regions. In contrast, the PRRX1-PAX3 Gln50 and PITX2-PAX3 Lys50 chimeras each showed nearly complete nuclear import and a diffuse distribution in the nucleus (Fig. 5-6C; *second and fourth rows*). Neither protein co-localized with Hoechst-rich areas, suggesting the PAX3 recognition helix alone is insufficient to cause heterochromatic localization in the context of another HD. Moreover, residue 50 is not sufficient to establish proper subnuclear compartmentalization, as displayed by the weak nucleolar localization of the PITX2-PAX3 Lys50 chimera. Consistent with the DNA binding studies above, these results suggest that residues in helices 1 and 2 cooperate with the recognition helix to maintain the structural and functional integrity of the HD. Surprisingly, residue 50 appears to play a role in this cooperativity, since substitution for Gln and Lys in the PRRX1-PAX3 and PITX2-PAX3 chimeras, respectively, improved protein stability and nuclear import, in addition to altering DNA binding activity. Thus, it appears that compatibility between residue 50 and other amino acids in the HD is required for proper HD behavior. Accordingly, subtle alterations to the overall HD structure as a consequence of position 50 substitutions may lead to altered DNA binding properties and compartmentalization in the nucleus.

5.8 The PAX3 homeodomain exhibits unique chromatin interaction properties

The subnuclear localization pattern of the PAX3 HD suggests a steady-state association with chromatin. To examine the basis for this interaction and to address differences between the PAX3, PRRX1, and PITX2 HDs with regard to chromatin interaction, we performed an *in vitro* affinity chromatography 'pull down' assay. Pull downs were carried out with bulk rat liver chromatin and recombinant HDs immobilized on nickel-agarose resin (see Section 2.6 and Figure 2-1). As shown, each HD pulled down core histones, although an obvious difference in interaction with linker histone-associated chromatin was observed: while both the PRRX1 and PITX2 HDs associated with chromatin containing histone H1, the PAX3 HD did not (Fig. 5-7A). Analysis of the chromatin-associated DNA bound by each HD showed that, despite its ability to localize to chromatin-rich regions in the nucleus, the PAX3 HD consistently pulled down the least DNA of the three HDs (Fig. 5-7B). These results present a paradoxical situation in which the PAX3 HD localizes to chromatin-rich regions of the nucleus, yet interacts poorly with chromatin and chromatin-associated DNA *in vitro*. We speculate that the PAX3 HD may be interacting with a different DNA population than the PRRX1 and PITX2 HDs, as suggested by the intensity curves in Fig. 5-7B. If this population is under-represented in the bulk chromatin, this could account for the overall reduction in DNA bound by the PAX3 HD.

PAX3 possesses two DNA-binding domains, the PD and HD, which are known to functionally interact, and we have also shown that the influence of the HD on subnuclear localization is suppressed in the full length protein. We therefore examined whether the presence of the PAX3 PD affects HD-chromatin interaction. For this, we again

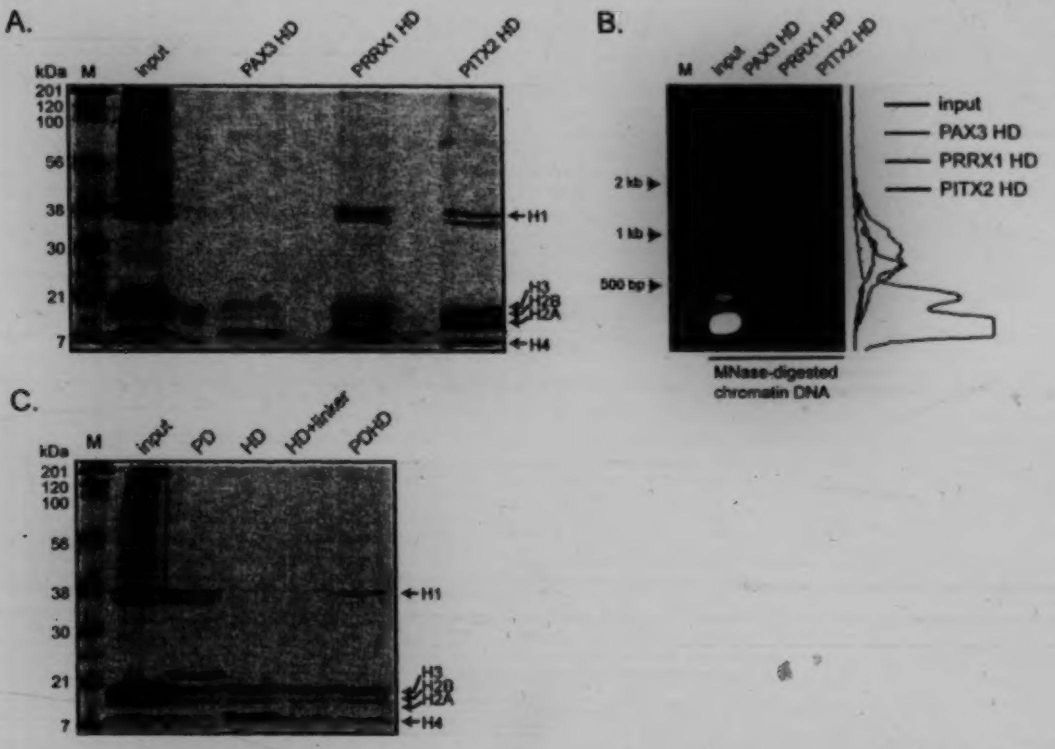

A.
kDa
M
input
PAX3 HD
PRRX1 HD
PITX2 HD
201
120
100
56
38
30
21
7
H1
H3
H2B
H2A
H4
B.
M
input
PAX3 HD
PRRX1 HD
PITX2 HD
2 kb
1 kb
500 bp
MNase-digested chromatin DNA
input
PAX3 HD
PRRX1 HD
PITX2 HD
C.
kDa
M
input
PD
HD
HD+linker
PDHD
201
120
100
56
38
30
21
7
H1
H3
H2B
H2A
H4

Figure 5-7. The PAX3 homeodomain possesses unique chromatin-interaction determinants. (A) Coomassie blue-stained gel used to resolve chromatin proteins that associate with the PAX3, PRRX1, or PITX2 HDs. Bulk rat liver chromatin was combined with immobilized PAX3 proteins, washed, and mixtures were resolved using denaturing polyacrylamide gel electrophoresis (PAGE) (see Section 2.6). Positions of core (H2A, H2B, H3, H4) and linker (H1) histone proteins are indicated. M, protein molecular weight marker (masses in kDa are indicated to the left of the gel). (B) Ethidium bromide-stained agarose gel used to resolve chromatin-associated DNA bound to the PAX3, PRRX1, and PITX2 HDs. Bulk rat liver chromatin was digested with MNase (0.25 U/100 μL chromatin), mixed with each immobilized HD, washed, stripped of protein, and resolved on a 0.7% agarose gel (see Section 2.6). M, DNA size marker (bp). Intensity curves of the material in each lane are shown at the right; line segments of each lane were plotted in Microsoft Excel and are represented graphically. (C) Coomassie blue-stained gel used to resolve chromatin proteins that associate with recombinant PAX3 domains. Pull-down assays were performed as described above. Positions of the core and linker histones are indicated. M, protein molecular weight marker (kDa).

performed an affinity chromatography pull down using bulk chromatin and recombinant portions of PAX3 containing the PD, HD, or PDHD, and an additional construct containing the HD and interdomain linker (HD+linker). We found that each PAX3 DNA-binding domain associated with chromatin and that the presence of the interdomain linker or physical linkage of both domains did not appreciably affect the affinity of these proteins for core histones (Fig. 5-7C). However, the PD interacted with linker histone-associated chromatin, and tethering the PD to the HD resulted in association with chromatin containing linker histone, indicating that the presence of the PD modulates chromatin interaction by the HD. These results support existing evidence (Underhill et al. 1995; Fortin et al. 1997; Underhill and Gros 1997; Apuzzo and Gros 2002; Apuzzo et al. 2004; Corry and Underhill 2005) which suggests that, in the context of full length PAX3, the PD (and possibly other regions of the protein) constrains HD behavior, and in the absence of these constraints, the HD adopts distinct localization and chromatin-interaction properties.

5.9 Discussion

In this chapter, we have investigated the behavior of the PAX3 HD and two related Prd-type HDs, both *in vitro* and in the context of the nucleus. We demonstrate that the PAX3, PRRX1, and PITX2 HDs display unique subnuclear compartmentalization patterns despite comparable intranuclear mobility and overlapping target preference. We also provide evidence that compatibility between amino acids in the HD recognition helix and those in helices 1 and 2 plays an important role in HD DNA binding and subnuclear localization. Specifically, substitution at residue 50 appears to induce subtle changes in the conformation of the HD, affecting its ability to be properly localized in the cell. Together, these observations suggest the presence of multiple determinants that modulate HD behavior.

The distinct intranuclear distributions of the PRRX1 and PITX2 HDs indicate that the PAX3 HD features specific determinants for localization to condensed chromatin. It is notable that these HDs exhibit only slight differences in their preference for 5'-TAATNN DNA recognition sites (Fig. 5-3), suggesting that DNA binding specificity does not dictate their subnuclear localization. In agreement, we have shown that no HD target motifs are preferentially enriched in the mouse genome, and that the preferred targets of the PAX3, PRRX1, or PITX2 HD show no regional enrichment on a representative mouse chromosome (Figs. 5-4D & E). Compartmentalization of HD proteins may therefore provide an additional level of specificity that limits their access to overlapping target sites or co-regulatory proteins. This idea is supported by studies of the MSX1 homeoprotein demonstrating that its compartmentalization and access to target genes is regulated by interaction with the PIAS1 protein (Lee et al. 2006). Our results also support current evidence suggesting that HD target specificity *in vitro* does not always correlate with functional specificity *in vivo*. Other attempts to reconcile target preference *in vitro* and functional specificity *in vivo* find that the two do not necessarily correlate, as we have shown here. For example, Bcd, a Lys50 homeoprotein, can activate *in vivo* reporter genes carrying appropriate reiterated binding sites (Driever et al. 1989b), while the fushi tarazu (Ftz) homeoprotein, which contains a Gln at position 50, cannot (Hanes et al. 1994). Moreover, a Gln50 version of Bcd retains this ability when the reporter gene has been modified to contain Gln50-optimized binding sites (Schier and

Gehring 1992; Hanes et al. 1994). In the case of Bcd, this activity is conferred entirely by the HD, as proteins containing the HD fused to a heterologous transactivation domain can activate reporters *in vivo* and rescue Bcd mutants (Driever et al. 1989a). Similar to our findings here, these results suggest inherent differences in the ability of homeobox proteins to recognize their cognate sites *in vivo* that were not apparent from their *in vitro* characterization.

The defining feature of the HD is its HTH structure, which allows the HD to interact with DNA and simultaneously project side chains from residues in helices 1 and 2 into its surroundings. Here, we have provided evidence suggesting an element of compatibility or "cross-talk" between amino acids in helices 1 and 2 and the recognition helix, in particular at position 50, is required for normal HD behavior. Although our results do not directly indicate that mutating residue 50 interferes with the folding or conformation of the HD, the fact that substituting Ser50 for Gln or Lys permitted efficient nuclear import and subnuclear localization of our chimeric HDs (Fig. 5-5C) implies that the conformation of these HDs was sufficiently stabilized upon substitution of residue 50. In addition, compatibility between the recognition helix and the rest of the HD is thought to influence target recognition by allowing for optimal positioning on a DNA template. In this regard, residues in helices 1 and 2 would play an important role in HD DNA binding behavior and numerous studies have shown that residues in this region contribute to binding affinity and specificity. For example, analyses of chimeric proteins containing amino acid substitutions in the Ftz HD showed that positions not involved in contacting DNA bases are important in HD target recognition *in vitro* and *in vivo* (Furukubo-Tokunaga et al. 1992), while mutational studies of the MSX1 HD showed that residues in helices 1 and 2 affected DNA binding behavior (Isaac et al. 1995). In addition, studies of the En HD showed that in the wild type protein, Gln50 exerted only partial influence on target specificity, implicating other HD residues in this behavior (Ades and Sauer 1994). Furthermore, positioning of residues in the amino-terminal extension of the DNA-bound En Q50K HD (Tucker-Kellogg et al. 1997) are altered compared to those in the wild type En HD (Ades and Sauer 1995), suggesting a conformational change induced by substitution at position 50. Finally, alanine scanning mutagenesis of the En HD across residues 17 to 46 showed that mutation of certain residues in this region have both positive and negative effects on DNA binding behavior (Sato et al. 2004). Together with the fact that human diseases are associated with mutations throughout the HDs of numerous homeobox transcription factors (D'Elia et al. 2001), these observations strongly imply that amino acids that do not directly contribute to target specificity or preference nevertheless make an important contribution to HD DNA binding properties.

In addition to the effects on HD DNA binding and overall structural integrity described above, our results also provide strong evidence that amino acids outside of the recognition helix influence HD behavior *in vivo*. Given the extensive amino acid identity within the recognition helix of the PAX3, PRRX1, and PITX2 HDs (with the exception of residue 50), these determinants are likely confined to helices 1 and 2. Most notably, there are dramatic differences in the charge properties of these HDs' solvent-exposed residues, creating a hierarchy among the three HDs in which PRRX1 possesses the highest isoelectric point and PAX3 the lowest (Fig. 5-1B). Subtle differences between solvent-exposed residues may allow regulatory proteins, for example, PBX (Chang et al.

1995) and MEIS (Chang et al. 1997) in the case of HOX proteins, and PIT1 in the case of PITX2 (Amendt et al. 1998), to discriminate among distinct HD surfaces, thereby influencing interaction with DNA. Although these interactions are thought to enhance HD affinity for target sites of increased complexity, they are sustained by contact with the HD solvent-exposed surface. In accordance, we and others have shown that, even in the absence of other proteins, solvent-exposed arginines play important roles in maintaining proper DNA binding behavior of the PRRX1 (Grueneberg et al. 1995) and PAX3 (GNC & DAU, unpublished data) HDs. Thus, it appears likely that amino acids outside of the recognition helix exert significant influence on all aspects of HD behavior, including DNA binding affinity, target preference, and subnuclear compartmentalization.

Finally, our results demonstrate that the PAX3 HD exhibits unique behavior in the presence of chromatinized templates. The observation that the PAX3 HD associated with a lower amount of chromatinized DNA than the PRRX1 or PITX2 HDs (Fig. 5-6B) suggests that structural impediments, be they core histones in the case of nucleosomal DNA or linker histone in the case of internucleosomal linker DNA, may restrict the ability of the PAX3 HD to access these templates. Given the promiscuous DNA binding behavior and relaxed target preference of HDs *in vitro* (Fig. 5-3), restriction of access to DNA templates may act as a valuable control mechanism *in vivo* that limits spurious binding of HDs to potential recognition sites. Alternatively, the PAX3 HD may possess a unique ability to displace linker histones from chromatin-embedded targets, presenting an interesting possibility that relates to PAX3 function during myogenesis. PAX3 lies genetically upstream of *MyoD* and is involved in *MyoD* expression (Maroto et al. 1997; Tajbakhsh et al. 1997; see Section 1.1.6), while the MSX1 homeoprotein plays an antagonistic role, inhibiting PAX3 DNA binding and suppressing *MyoD* expression (Bendall et al. 1999). It was shown that MSX1 cooperates with histone H1b to inhibit the myogenic pathway by creating a repressive chromatin structure at *MyoD* core enhancer region (Lee et al. 2004). Although the relationship between the PAX3 HD and linker histones requires further investigation, it is tempting to speculate that an aspect of *MyoD* activation involves displacement of linker histones from the *MyoD* enhancer region by PAX3. Interestingly, the regions of MSX1 responsible for histone H1b interaction and *MyoD* repression include the HD and a short sequence amino-terminal to the HD featuring an Eh-1-like motif (Lee et al. 2004). By analogy, the HD (and possibly the octapeptide) of PAX3 may be responsible for associating with histone H1 and facilitating its removal from chromatin templates in certain situations.

CHAPTER 6: CONCLUSIONS & OUTLOOK

6.1 Summary

The PAX3 transcription factor plays key roles in metazoan development and mutations in the protein cause the *Sp* phenotype and WS in mice and humans, respectively. Through molecular and cellular analyses, we have investigated the determinants behind PAX3 behavior, including DNA binding, gene transactivation, subnuclear localization, and intranuclear mobility. We show that differential use of the PAX3 PD and HD permits interaction with a variety of recognition sequences through distinct modes of DNA binding and that functional cooperativity between the two domains underlies this behavior. Our investigations of PAX3 in a cellular context show that the protein is localized in the interchromatin compartment, displaying limited co-localization with sites of transcriptional activity. Furthermore, we demonstrate that PAX3 possesses multiple determinants that regulate compartmentalization and mobility in the nucleus. We have also shed new light on the properties of the PAX3 HD, both in the context of the full length protein and in comparison to other Prd-type HDs. Our most significant findings, however, demonstrate that disease mutations exert a wide range of effects on PAX3 behavior. We have shown that mutations associated with WS and the Sp^d phenotype can have severe effects on PAX3 DNA binding and inter-domain cooperativity, and that these effects do not necessarily correlate with trans-regulation of PAX3 target genes. Moreover, these mutations exerted pleiotropic effects on the subcellular localization of PAX3, although all the mutants tested exhibited altered intranuclear mobility, reflecting a fundamental defect in the molecular interactions that control PAX3 function *in vivo*. In summary, the observations presented here provide important information regarding the determinants that regulate PAX3 behavior and the effects of disease mutations on these regulatory mechanisms.

6.2 A two-step model for regulation of PAX3 function *in vivo*

A key to understanding the function of transcriptional regulators involves characterizing the determinants responsible for directing them to target sites in the nucleus. We know that DNA-binding proteins must efficiently recognize and interact with chromatin-embedded target sites, but currently there is a poor understanding of how this occurs *in vivo*. The dynamic nature of transcription factors and chromatin templates suggests that a great deal of coordination is involved in localizing a regulatory DNA sequence and its cognate binding protein to the nuclear domain in which interaction is to occur. Since a DNA-binding protein must localize to this domain *and* interact with its target sequence, it becomes pertinent to ask whether subnuclear compartmentalization is a cause or effect of the functional behavior of transcription factors. Put another way, is compartmentalization of a transcription factor determined by the target loci it interacts with, or does localization to a particular compartment precede and therefore dictate the targets with which the protein interacts? With regard to PAX3, our investigations of disease mutants provide some insight into this question and evoke a two-step process that regulates PAX3 function *in vivo*. Among the PAX3 mutants studied here, there appears to be an inverse correlation between DNA binding ability and intranuclear compartmentalization, where class I mutants (N47H, G81A, V265F) showed the most severe localization defects despite relatively normal or enhanced DNA binding, while class II mutants (Sp^d, F45L, S84F, Y90H, R271H), which most closely resembled wild type PAX3 in the nucleus, possessed severe DNA binding defects. These trends were also evident in analyses of intranuclear mobility, where the PAX3-GFP F45L/G81A protein resembled a class I mutant in live cells. Together, these results show that failure to properly compartmentalize precludes any chance of target interaction, regardless of DNA binding capability, and strongly suggest that localization to a particular nuclear region is the first step in PAX3 target gene regulation.

Once compartmentalized, PAX3 must then locate and bind to a target regulatory sequence. The number of genes PAX3 regulates is currently unknown, but they probably represent a very small proportion of all potential recognition sites available to bind at any time. The mechanism that controls how PAX3 identifies and interacts with its target binding sites therefore constitutes the second step in regulating PAX3 function in the nucleus and also represents the underlying defect of class II mutations. This mechanism likely depends on a number of factors, including local chromatin structure, interactions with nearby protein complexes, such as the basal transcriptional machinery or accessory binding proteins, and, ultimately, the DNA sequence. Finally, once its localization at a particular target is established, PAX3 would activate expression of the gene, possibly through the recruitment of chromatin-modifying enzymes. Related to this, it was recently shown that PAX7 interacts with the WDR5-ASH2L-MLL2 complex, which methylates Lys4 of histone H3, and recruits PAX7 to the *Myf5* promoter (McKinnell et al. 2008). Given its homology to PAX7, we might expect PAX3 to also recruit H3K4-methylating machinery to target loci, which could also account for the regions of co-localization we observed between the two signals (Fig. 4-2A, C-E). Thus, our demonstration that the two mutational classes partition according to the severity of effects on (a) intranuclear

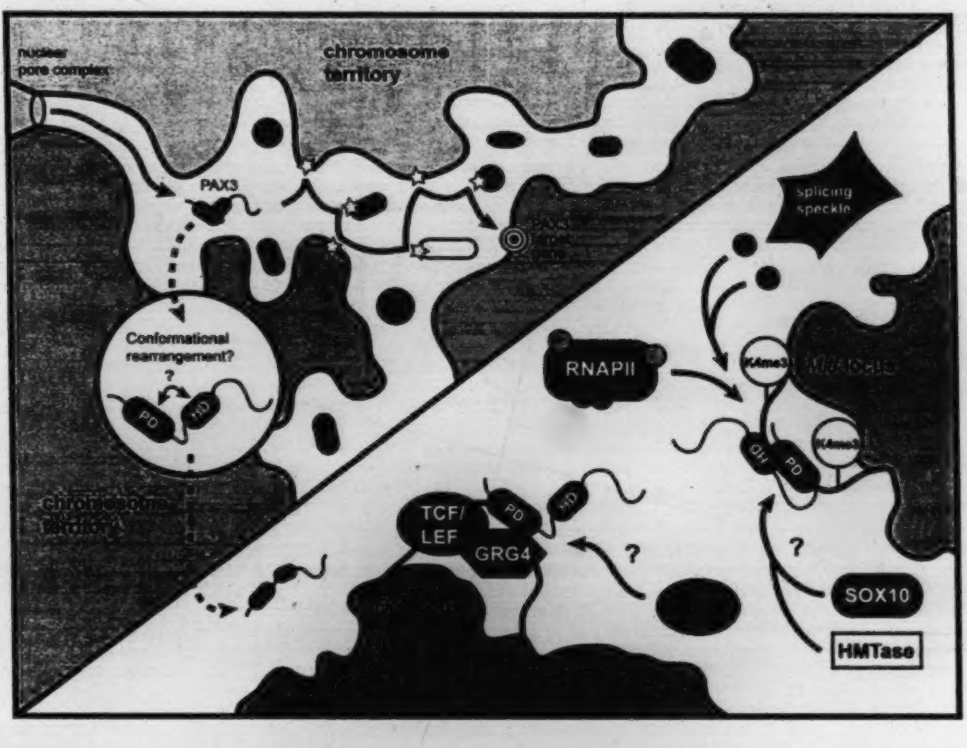


Figure 6-1. A two-step model for regulation of PAX3 function *in vivo*. *Top*, upon import into the nucleus, PAX3 is compartmentalized in the interchromatin compartment (grey). PAX3 compartmentalization may be dependent on the adoption of a specific conformation(s), which would suppress determinants of chromatin localization. Class I disease mutations interfere with the nuclear import and/or retention and intranuclear compartmentalization of PAX3, suggesting a disturbance to this conformation. PAX3 is mobile in the interchromatin compartment and is likely targeted to chromatin-embedded recognition sites through random, stochastic interactions (represented by yellow stars) with nuclear components. Upon reaching a target locus, PAX3 may undergo a conformational rearrangement (*inset*, *left*), that permits interaction with regulatory DNA sequences; this may be achieved in association with molecular chaperones. *Bottom*, PAX3 targets are likely located on chromatin fibers looped from chromosome territories into the interchromatin space. Target promoter sequences and use of the paired domain (PD; blue) and homeodomain (HD, red) regulate how PAX3 interacts with these targets; examples shown are *Mitf*, which is activated by PAX3, and *Dct*, which is repressed by PAX3 (see Sections 1.1.6 and 1.1.7). Note the differential use of the PD and HD in binding to the two sites, which is inferred from our DNA binding data. Subsequent to or concomitant with PAX3 binding at these loci, components of the transcriptional machinery, including RNA polymerase II (RNAPII), splicing and chromatin remodeling factors (e.g., histone methyltransferases (HMTase)), and possibly accessory transcription factors (e.g., SOX10) in the case of *Mitf*, and co-repressors and possibly histone deacetylases (HDAC) in the case of *Dct*, are recruited. Class II disease mutations are thought to disrupt molecular interactions at this stage, permitting compartmentalization but exerting variable defects in mobility and DNA binding activity.

compartmentalization (class I) and (b) DNA binding (class II) supports a two-step process for regulating PAX3 function in the nucleus (summarized in Fig. 6-1). Significantly, this model can be applied to other transcription factors whose function is dependent on compartmentalization, such as runt family proteins, which require localization to the nuclear matrix for proper transcriptional regulatory activity (see Section 1.3.3).

6.3 Do distinct conformations underlie PAX3 behavior and function?

Our investigations demonstrate that class I mutations do not cluster to a specific region of the protein, implying that they do not affect any one domain or substructure within PAX3, but instead perturb its overall conformation. Without structural data, it is difficult to state with certainty that PAX3 adopts a certain conformation(s) *in vivo*, and whether disease mutations interfere with protein folding or cause a novel conformation to be assumed; however, several points of data suggest that distinct conformational arrangements of PAX3 play an important role in optimizing its function. First, there is extensive evidence that functional cooperativity between the PAX3 PD and HD is required for DNA binding activity (Underhill et al. 1995; Fortin et al. 1997; Underhill and Gros 1997; Corry and Underhill 2005a). Scanning mutagenesis and computer modeling indicate that the PD and HD would be juxtaposed upon binding to DNA (Apuzzo and Gros 2007), reflecting the proposed arrangement of the Prd PD and HD bound to the *PH0* site (Jun and Desplan 1996). In these models, the HD would likely be "looped back" to interact with the DNA immediately upstream of the PD-binding site, which would also account for both the conserved arrangement of the PD and HD motifs in the *MITF* and *MYF5* promoter sites and the high affinity of PAX3 for these sequences. Secondly, our results show that the acidic linker region (ALR) between the PD and octapeptide appears to exert a negative effect on HD DNA binding activity and interdomain cooperativity, and that other regions of PAX3 appear to mask or obscure this region, alleviating its inhibitory action (see Appendix 1). In considering a model where the HD "loops back" to bind adjacent to the PD, the ALR would be brought into proximity to other parts of PAX3, including the octapeptide and a stretch of regularly spaced serine residues between the octapeptide and HD. Significantly, this serine-rich region resembles a flexible domain in ETS1 that contains several phosphoserines and is responsible for allosterically controlling the DNA binding activity of the protein (Pufall et al. 2005). Together with the high probability that the serines of the PAX3 linker are phosphorylated (data not shown), our data suggest that the head-to-head conformation of the DNA-bound PD and HD may bring the ALR and polyserine regions close together, possibly suppressing the autoinhibitory effects of the former (Fig. 6-2). Alternatively, the PAX3 PD and HD could interact directly, as shown for the PAX6 PD and various Prd-type HDs (Bruun et al. 2005). In this scenario, residues in the HD form an interface with residues of the carboxy-terminal PD subdomain, bringing the domains in contact in a DNA-independent manner. Although this type of interaction has yet to be demonstrated for the PAX3 PD or HD, a conformation such as this also has the potential to mask the ALR, and could also act to constrain the HD, suppressing its effects on intranuclear compartmentalization until it is required for DNA binding (see below). Future experiments, such as mutating residues in the ALR or swapping the linkers between

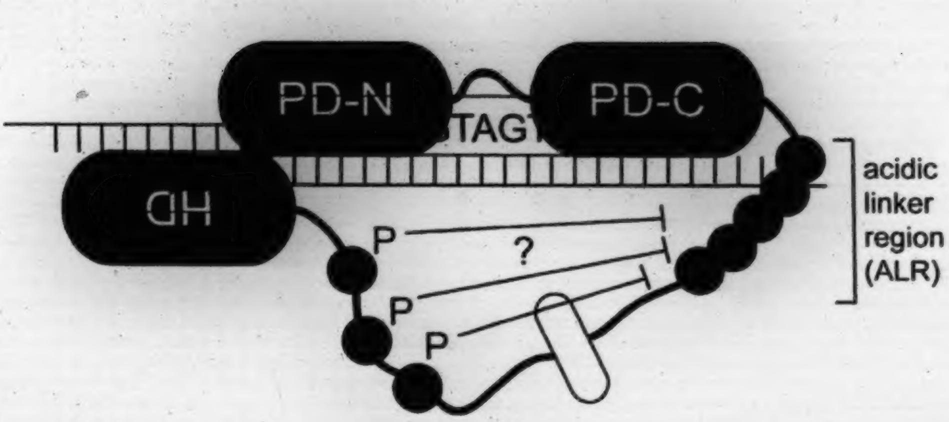


Figure 6-2. A putative composite sequence-binding conformation may suppress an autoinhibitory region in the PAX3 linker. In PAX3, the interdomain linker sequence that immediately follows the paired domain (PD, blue), termed the acidic linker region (ALR; green), appears to exert an inhibitory effect on the homeodomain (HD, red) (see Appendix 2). DNA binding studies with the *MITF/Mitf* promoter sequence (over which PAX3 is superimposed) suggest that the HD binds to a motif upstream of the PD, which may induce a conformation that suppresses the inhibitory effect of the ALR. Alternatively, this conformation could allow a series of phosphoserines (orange) following the octapeptide (yellow) to mask the ALR (see Section 6.4 for details).

PAX3 and other PAX factors, should more clearly resolve the determinants of the linker that regulate PAX3 function.

In vivo evidence for a specific conformational arrangement of PAX3 comes from comparing the subnuclear localization of the full length protein and the isolated HD. The influence of the HD on PAX3 compartmentalization is obviously suppressed in the context of the full length protein, suggesting that PAX3 is folded in such a way so as to conceal determinants in the HD that promote interaction with chromatin. However, we and others have demonstrated that the HD plays an important role in PAX3 target recognition and affinity, implying that the domain would have to be freed from occlusion once PAX3 is in proximity of potential binding sequences. It is therefore interesting to imagine that upon nuclear import, PAX3 exists in one conformation that simultaneously exposes the determinants that mediate localization to the interchromatin compartment and obscures those, such as the HD, that would direct PAX3 to other nuclear domains, for example, pericentromeric heterochromatin. Once PAX3 reaches its target locus, the protein would then undergo a conformational change, releasing the HD to participate in DNA binding or protein-protein interactions (see Fig. 6-1). Indeed, structural "fuzziness," or the inherent polymorphism of proteins in their native state, has been shown to be a key feature of protein-protein interactions (Tompa and Fuxreiter 2008), and may also apply to protein-DNA interactions, as we propose here. In this scenario, chaperone proteins might be involved in maintaining and stabilizing DNA-bound and/or unbound PAX3 conformations, as has been shown for other transcription factors (Hager et al. 2006).

Finally, support for specific conformational arrangements of PAX3 is provided by our observations of disease-associated mutants. Class I mutations impair nuclear import or retention, suggesting that the structure of PAX3 may be altered so as to partially obscure the NLS or possibly expose a nuclear export signal, but support DNA binding activity, including PD-HD cooperativity (see Figs. 3-11 & 3-12). Conversely, class II mutants display compartmentalization patterns resembling that of wild type PAX3, suggesting little interference with folding or structural stability. However, the increased mobility of these mutants relative to the wild type protein suggests a defect in the determinants that constrain the movement of PAX3 in the nucleus, such as interactions with target binding sites (see Section 4.9), as reflected in the impairments to DNA binding (e.g., F45L, S84F) and PD-HD cooperativity (e.g., R271G) observed in our *in vitro* assays. Thus, if we imagine that PAX3 adopts two conformations *in vivo*, one that facilitates correct subnuclear compartmentalization and one that supports DNA binding activity, our mutational classes can also be segregated according to the conformation they disturb. These observations are also consistent with the proposed two-step model of PAX3 regulation (above), in which subnuclear compartmentalization (mediated by one conformation) brings PAX3 into the vicinity of target loci, followed by interaction with specific target regulatory sequences (supported by a second, distinct conformation).

6.4 PAX3: activator *and* repressor?

The PAX proteins are generally thought of as transcriptional activators, yet multiple studies provide evidence that PAX3 also acts as a repressor of transcriptional activity (Chalepakis et al. 1994b; Kallunki et al. 1995; Kioussi et al. 1995; Tremblay et al. 1996; Kwang et al. 2002; Lang et al. 2005). Interactions between PAX3 and proteins such as HIRA (Magnaghi et al. 1998) and pRB family members (Wiggan et al. 1998) also suggest a possible role for PAX3 in cellular senescence or in maintaining some loci in a permanently repressed state. Additionally, Gro family proteins, which include the PAX3-interacting GRG4 co-repressor, mediate locus-specific repression, but are also known to participate in long-term gene silencing, possibly through targeted spreading of deacetylated chromatin domains or formation of a repressosome that interacts with the basal transcription machinery (reviewed in Courey and Jia 2001). Lastly, studies comparing gene regulation by PAX3 and PAX3-FKHR have revealed a subset of targets that are repressed by the former but activated by the latter (Begum et al. 2005). Thus, it appears that PAX3 may play a role both in gene activation and repression, depending on the context. However, the mechanism underlying the switch between activation and repression remains unclear.

The regulation of melanogenesis by PAX3 (see Section 1.1.6) provides a well-characterized model of its dual activation/repression activity. PAX3 appears to be able to simultaneously activate *Mitf* expression, promoting the maintenance of melanogenic precursors, and repress genes involved in terminal differentiation (e.g., *Dct*) in the same cell population (Lang et al. 2005). This situation is reminiscent of the dual role of the dorsal (Dl) protein as a transcriptional activator and repressor during dorso-ventral patterning of the *Drosophila* embryo (Ray et al. 1991). Dl interacts with Gro through an Eh-1-like motif, facilitating repression of dorsal fate genes (Dubnicoff et al. 1997; Flores-Saaib et al. 2001), and Dl-mediated repression is enhanced by the concomitant binding of the *cut* and *dead ringer* gene products to AT-rich sites in ventral repression regions

(Valentine et al. 1998). Similarly, during the development of *Drosophila* cone cells, the runt domain transcription factor lozenge (Lz) is converted to a repressor upon interaction with Gro and the Cut protein, which binds to AT-rich sequences that flank the Lz-binding element (Canon and Banerjee 2003). It is therefore possible that binding site context may also determine whether PAX3 behaves as an activator or repressor in a given cell. In this regard, PAX3 binding to low affinity PD-only recognition sequences (e.g., *Dct*) versus high affinity composite sites (e.g., *MITF*) may determine the transcriptional output at a given locus. As such, PAX3 interaction with low affinity sites may require coincident binding of other factors to achieve full activity, as shown for *Dct*, where TCF/LEF binding was necessary for PAX3-mediated repression (Lang et al. 2005).

Differences in binding site composition may also induce allosteric effects that determine activation or repression, particularly in the case of a multi-domain transcription factor like PAX3. The allosteric control of transcriptional regulators by target binding sites is a well-known method of modulating gene expression (Lefstin and Yamamoto 1998) and the regulation of PIT1 target gene expression clearly demonstrates this potential. PIT1 binds to conserved motifs in the growth hormone (GH) gene promoter, activating GH expression in somatotropes and repressing expression in lactotropes (Lira et al. 1988). In lactotropes, a 2-bp insertion in one of these motifs induces an altered PIT1 conformation that supports interaction with the N-CoR co-repressor, leading to repression of *GH*, while in somatotropes, the *GH* promoter lacks this insert, allowing PIT1 to activate GH expression (Scully et al. 2000). In the case of PAX3, we predict that its DNA-bound conformation might differ in the presence of PD-only versus composite binding sites, potentially producing different interaction surfaces for partner proteins or modulating affinity for chromatin-embedded targets. Structural studies of a minimal PAX3 protein, for example the PDHD, bound to representatives of the two target types will likely provide the definitive answer to how PAX3 is positioned on DNA targets and how the PD and HD contribute to different binding modes.

A final aspect of establishing the repressive qualities of PAX3 involves examining disease-causing mutants. To date, studies of disease mutations have provided little insight into the repressive role of PAX3. Using the PAX3-responsive *MITF-Luc* and *Trp1-Luc* constructs, we found that, in general, class I mutants either had neutral or repressive effects on reporter gene activation, while class II mutants tended to activate expression of the reporter genes (Fig. 3-13). The outlier is PAX3 R271G, which significantly repressed both *MITF-Luc* and *Trp1-Luc* below baseline levels. PAX3 G81A was the only other mutant to cause this level of repression. Significantly, these were the only mutations in our cohort to consistently mislocalize PAX3 to TLE4b and TLE4c foci (see Appendix 2). Given that TLE4b and TLE4c foci resemble matrix-associated deacetylase (MAD) repression bodies formed by other co-repressors (see Downes et al. 2000), it is possible that recruitment of PAX3 G81A and R271G to these structures endows them with repressive activity through their contact with TLE4, histone deacetylases (HDACs), and/or other components of the repression machinery. This parallels the scenario in which the hypermethylated in cancer-1 (HIC1) protein recruits the TCF4 transcription factor to HIC1 nuclear bodies, causing repression of TCF4 target gene activation in response to Wnt signaling (Valenta et al. 2006). Interestingly, the mechanism that underlies the recruitment of PAX3 to TLE4 foci appears to dominate over DNA binding ability and subnuclear localization, since these properties are retained by the G81A and R271G mutants, respectively. Thus, it will be interesting to determine what causes PAX3 to lose its ability to redistribute these TLE4 isoforms and what triggers association between the two proteins in focal domains.

6.5 PAX3 and Waardenburg syndrome: towards a genotype-phenotype correlation

To date, no obvious genotype-phenotype correlations have been identified among currently described PAX3 mutations. This is exemplified by the observation that mutation of Asn47 to basic residues causes two clinically distinct disorders – Asn47→Lys was found in a family with CDHS (Asher et al. 1996b), while mutation to histidine was found to cause WS3 (Hoth et al. 1993). A previous study that attempted to examine WS genotype-phenotype correlations suggested that patients with deletions of the HD or Pro-Ser-Thr-rich domain have increased susceptibility to pigmentary disturbances compared to patients with missense mutations in the HD (DeStefano et al. 1998). Additionally, a Pro→Leu mutation that characterizes a Brazilian family (da-Silva 1991; Baldwin et al. 1992) appears to be associated with hearing loss, since a large proportion (>78%, compared to a ~20%-30% average in WS1 families) of WS1 individuals in this kindred exhibit sensorineural deafness. Nevertheless, the lack of a definitive phenotype-genotype relationship points to other factors, including stochastic changes in PAX3 expression or imbalances in enhancer activity, which may influence a particular phenotype in a cell-type specific manner.

Although our results provide no clear genotype-phenotype correlations among the cohort of mutants examined here, some broad trends emerge upon integration of the data. In general, class I mutants exhibit near-normal DNA binding activity, including PD-HD cooperativity, have neutral or repressive effects on reporter gene activation, and exhibit a defect in subnuclear compartmentalization and mobility. In contrast, class II mutants display problems in DNA binding, including reduced affinity and impaired PD-HD cooperativity, tend to activate reporter gene expression (with the exception of R271G, as discussed above), and exhibit minor effects on intranuclear compartmentalization and mobility. Paradoxically, the N47H and G81A mutations are associated with the most and least severe manifestations of WS, respectively, in the existing literature. Members of the family in which the G81A mutation was first identified show a relatively mild phenotype and were originally diagnosed with WS2 (Tassabehji et al. 1993), while the N47H mutation causes WS3 in the heterozygous state (Hoth et al. 1993). Together, these observations serve to emphasize the difficulty in establishing a genotype-phenotype correlation in patients carrying a *PAX3* mutation (especially considering these are both class I mutations). Furthermore, our data illustrate the problems in trying to assign a phenotype or diagnosis based on *in vitro* studies of mutant transcription factors, as has historically been the case. In the case of PAX3, it is evident that mutants possessing apparently normal DNA binding activity and/or promoter occupancy do not necessarily exhibit normal activity when assessed in a cellular context. Although more mutations should be examined to obtain a larger sample, our results provide a strong starting point for future analyses of how PAX3 mutations affect the protein's behavior.

6.6 Outlook

This is the first study to integrate *in vivo*- and *in vitro*-derived observations of the properties and behavior of a PAX protein, and provides important and useful information on how a developmental transcription factor is regulated, particularly in the context of the nucleus. We have also demonstrated several correlations between DNA binding activity, functional regulation of target gene activation, and intranuclear compartmentalization, and our assessments of how disease mutations affect these processes provide insight into the determinants that regulate PAX3 behavior *in vivo*. We anticipate that the results presented here can be applied to understanding how other PAX factors are regulated and how disease mutations affect their function, both at a molecular level and in a nuclear context. Given their roles in controlling key developmental processes, we might expect that the overall functional behavior of PAX proteins is similar, i.e., each interacts with the regulatory elements of target genes to activate or repress that gene's expression. However, differences exist among the PAX proteins, both obvious (the absence or presence of the HD and/or octapeptide) and subtle (amino acid composition of DNA binding and transactivation domains). Effects of the latter class of differences are evident in our recruitment assays using PD and HD paralogs (Figs. 3-9 & 3-10) and, importantly, provide a foundation for investigating the intrinsic determinants that specify the behavior of the nine PAX proteins.

This study also provides important information regarding the general understanding of transcription factor function. Firstly, our analysis of PAX3 binding sequences demonstrates that *in vivo* targets may not necessarily resemble *in vitro*-derived optimal target sequences. For a multi-domain factor like PAX3, it must be considered that a binding site may contain recognition motifs for one or more of its DNA-binding domains. Furthermore, the evolutionary conservation of a promoter element must be analyzed, which should indicate whether or not the sequence is a bona fide functional target. Lately, ChIP-chip, or ChIP followed by hybridization to microarray chips, has emerged as a powerful method to examine the *in vivo* occupancy of DNA by transcription factors (Hanlon and Lieb 2004). Parallel ChIP-chip experiments to investigate chromatin modifications at these sites allow researchers to create a more detailed picture of the local environment in which a transcription factor functions. Additionally, the use of microarray chips representing the gene expression profiles of specific tissues or cellular processes provides an opportunity to investigate the activity of a transcription factor in the context of a range of physiological conditions and systems. Thus, ChIP-chip should prove useful in clarifying PAX3 target regulation during specific developmental stages and pathways, as well as addressing its postnatal functions and involvement in cancer and disease. In addition, a novel ChIP-sequencing method, which bypasses the requirement for microarray chips, has recently been used to identify *in vivo* binding sites for the neuron-restrictive silencer factor (Johnson et al. 2007). Assays such as this should facilitate the identification of transcription factor interaction sites across entire genomes and allow a more precise description of target motifs.

Finally, we and others have demonstrated that disease mutations can exert a range of effects on the functional properties of transcription factors. With improvements in the techniques used to examine protein behavior *in vivo*, it is becoming clear that the effects mutations have on a protein's behavior *in vitro* are not always a reliable indicator of how its behavior will be affected at the cellular level. For example, each mutation we analyzed affected PAX3 mobility in the nucleus, despite the range of effects they had on PAX3 behavior in DNA binding and transactivation assays. Results of our live cell experiments, together with our DNA binding analyses, also provide important information on how the intranuclear compartmentalization and functional behavior of transcription factors are related and how their underlying determinants control transcription factor activity *in vivo*. Finally, we anticipate that the observations made here will enhance our understanding of PAX3 behavior and clarify how mutations in the protein contribute to the onset and progression of PAX3-related disorders.

BIBLIOGRAPHY

Ades, S.E. and R.T. Sauer. 1994. Differential DNA-binding specificity of the engrailed homeodomain: the role of residue 50. *Biochemistry* **33**: 9187-94.

-. 1995. Specificity of minor-groove and major-groove interactions in a homeodomain-DNA complex. *Biochemistry* **34**: 14601-8.

Agresti, A., P. Scaffidi, A. Riva, V.R. Caiolfa, and M.E. Bianchi. 2005. GR and HMGB1 interact only within chromatin and influence each other's residence time. *Mol Cell* **18**: 109-21.

Ahmad, K. and S. Henikoff. 2001. Modulation of a transcription factor counteracts heterochromatic gene silencing in Drosophila. *Cell* **104**: 839-47.

Akhtar, A. and S.M. Gasser. 2007. The nuclear envelope and transcriptional control. *Nat Rev Genet* **8**: 507-17.

Alexander, R.B., G.L. Greene, and E.R. Barrack. 1987. Estrogen receptors in the nuclear matrix: direct demonstration using monoclonal antireceptor antibody. *Endocrinology* **120**: 1851-7.

Altschul, S.F., W. Gish, W. Miller, E.W. Myers, and D.J. Lipman. 1990. Basic local alignment search tool. *J Mol Biol* **215**: 403-10.

Amendt, B.A., L.B. Sutherland, E.V. Semina, and A.F. Russo. 1998. The molecular basis of Rieger syndrome. Analysis of Pitx2 homeodomain protein activities. *J Biol Chem* **273**: 20066-72.

Apuzzo, S., A. Abdelhakim, A.S. Fortin, and P. Gros. 2004. Cross-talk between the paired domain and the homeodomain of Pax3: DNA binding by each domain causes a structural change in the other domain, supporting interdependence for DNA Binding. *J Biol Chem* **279**: 33601-12.

Apuzzo, S. and P. Gros. 2002. Site-specific modification of single cysteine Pax3 mutants reveals reciprocal regulation of DNA binding activity of the paired and homeo domain. *Biochemistry* **41**: 12076-85.

-. 2006. The paired domain of Pax3 contains a putative homeodomain interaction pocket defined by cysteine scanning mutagenesis. *Biochemistry* **45**: 7154-61.

-. 2007. Cooperative interactions between the two DNA binding domains of Pax3: helix 2 of the paired domain is in the proximity of the amino terminus of the homeodomain. *Biochemistry* **46**: 2984-93.

Asher, J.H., Jr., R.W. Harrison, R. Morell, M.L. Carey, and T.B. Friedman. 1996a. Effects of Pax3 modifier genes on craniofacial morphology, pigmentation, and viability: a murine model of Waardenburg syndrome variation. *Genomics* **34**: 285-98.

Asher, J.H., Jr., A. Sommer, R. Morell, and T.B. Friedman. 1996b. Missense mutation in the paired domain of PAX3 causes craniofacial-deafness-hand syndrome. *Hum Mutat* **7**: 30-5.

Bajard, L., F. Relaix, M. Lagha, D. Rocancourt, P. Daubas, and M.E. Buckingham. 2006. A novel genetic hierarchy functions during hypaxial myogenesis: Pax3 directly activates Myf5 in muscle progenitor cells in the limb. *Genes Dev* **20**: 2450-64.

Balczarek, K.A., Z.C. Lai, and S. Kumar. 1997. Evolution of functional diversification of the paired box (Pax) DNA-binding domains. *Mol Biol Evol* **14**: 829-42.

Baldwin, C.T., C.F. Hoth, J.A. Amos, E.O. da-Silva, and A. Milunsky. 1992. An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. *Nature* **355**: 637-8.

Baldwin, C.T., C.F. Hoth, R.A. Macina, and A. Milunsky. 1995. Mutations in PAX3 that cause Waardenburg syndrome type I: ten new mutations and review of the literature. *Am J Med Genet* **58**: 115-22.

Balling, R., U. Deutsch, and P. Gruss. 1988. undulated, a mutation affecting the development of the mouse skeleton, has a point mutation in the paired box of Pax 1. *Cell* **55**: 531-5.

Bang, A.G., N. Papalopulu, M.D. Goulding, and C. Kintner. 1999. Expression of Pax-3 in the lateral neural plate is dependent on a Wnt-mediated signal from posterior nonaxial mesoderm. *Dev Biol* **212**: 366-80.

Bang, A.G., N. Papalopulu, C. Kintner, and M.D. Goulding. 1997. Expression of Pax-3 is initiated in the early neural plate by posteriorizing signals produced by the organizer and by posterior non-axial mesoderm. *Development* **124**: 2075-85.

Barber, T.D., M.C. Barber, T.E. Cloutier, and T.B. Friedman. 1999. PAX3 gene structure, alternative splicing and evolution. *Gene* **237**: 311-9.

Barberis, A., G. Superti-Furga, L. Vitelli, I. Kemler, and M. Busslinger. 1989. Developmental and tissue-specific regulation of a novel transcription factor of the sea urchin. *Genes Dev* **3**: 663-75.

Barr, F.G. 1997. Chromosomal translocations involving paired box transcription factors in human cancer. *Int J Biochem Cell Biol* **29**: 1449-61.

Barr, F.G., J.C. Fitzgerald, J.P. Ginsberg, M.L. Vanella, R.J. Davis, and J.L. Bennicelli. 1999. Predominant expression of alternative PAX3 and PAX7 forms in myogenic and neural tumor cell lines. *Cancer Res* **59**: 5443-8.

Barrack, E.R. 1983. The nuclear matrix of the prostate contains acceptor sites for androgen receptors. *Endocrinology* **113**: 430-2.

-. 1987. Steroid hormone receptor localization in the nuclear matrix: interaction with acceptor sites. *J Steroid Biochem* **27**: 115-21.

Barrack, E.R. and D.S. Coffey. 1980. The specific binding of estrogens and androgens to the nuclear matrix of sex hormone responsive tissues. *J Biol Chem* **255**: 7265-75.

Barsony, J., I. Renyi, and W. McKoy. 1997. Subcellular distribution of normal and mutant vitamin D receptors in living cells. Studies with a novel fluorescent ligand. *J Biol Chem* **272**: 5774-82.

Baumann, C.T., H. Ma, R. Wolford, J.C. Reyes, P. Maruvada, C. Lim, P.M. Yen, M.R. Stallcup, and G.L. Hager. 2001. The glucocorticoid receptor interacting protein 1 (GRIP1) localizes in discrete nuclear foci that associate with ND10 bodies and are enriched in components of the 26S proteasome. *Mol Endocrinol* **15**: 485-500.

Becker, M., C. Baumann, S. John, D.A. Walker, M. Vigneron, J.G. McNally, and G.L. Hager. 2002. Dynamic behavior of transcription factors on a natural promoter in living cells. *EMBO Rep* **3**: 1188-94.

Begum, S., N. Emani, A. Cheung, O. Wilkins, S. Der, and P.A. Hamel. 2005. Cell-type-specific regulation of distinct sets of gene targets by Pax3 and Pax3/FKHR. *Oncogene* **24**: 1860-72.

Bendall, A.J., J. Ding, G. Hu, M.M. Shen, and C. Abate-Shen. 1999. Msx1 antagonizes the myogenic activity of Pax3 in migrating limb muscle precursors. *Development* **126**: 4965-76.

Bennicelli, J.L., R.H. Edwards, and F.G. Barr. 1996. Mechanism for transcriptional gain of function resulting from chromosomal translocation in alveolar rhabdomyosarcoma. *Proc Natl Acad Sci U S A* **93**: 5455-9.

Bentley, D.L. 2005. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Curr Opin Cell Biol* **17**: 251-6.

Berezney, R., M.J. Mortillaro, H. Ma, X. Wei, and J. Samarabandu. 1995. The nuclear matrix: a structural milieu for genomic function. *Int Rev Cytol* **162A**: 1-65.

Berger, S.L. 2002. Histone modifications in transcriptional regulation. *Curr Opin Genet Dev* **12**: 142-8.

Berkes, C.A., D.A. Bergstrom, B.H. Penn, K.J. Seaver, P.S. Knoepfler, and S.J. Tapscott. 2004. Pbx marks genes for activation by MyoD indicating a role for a homeodomain protein in establishing myogenic potential. *Mol Cell* **14**: 465-77.

Bernasconi, M., A. Remppis, W.J. Fredericks, F.J. Rauscher, 3rd, and B.W. Schafer. 1996. Induction of apoptosis in rhabdomyosarcoma cells through down-regulation of PAX proteins. *Proc Natl Acad Sci U S A* **93**: 13164-9.

Bernstein, E. and C.D. Allis. 2005. RNA meets chromatin. *Genes Dev* **19**: 1635-55.

Berry, F.B., M.A. Lines, J.M. Oas, T. Footz, D.A. Underhill, P.J. Gage, and M.A. Walter. 2006. Functional interactions between FOXC1 and PITX2 underlie the sensitivity to FOXC1 gene dose in Axenfeld-Rieger syndrome and anterior segment dysgenesis. *Hum Mol Genet* **15**: 905-19.

Berry, F.B., M.A. O'Neill, M. Coca-Prados, and M.A. Walter. 2005. FOXC1 transcriptional regulatory activity is impaired by PBX1 in a filamin A-mediated manner. *Mol Cell Biol* **25**: 1415-24.

Bertuccioli, C., L. Fasano, S. Jun, S. Wang, G. Sheng, and C. Desplan. 1996. In vivo requirement for the paired domain and homeodomain of the paired segmentation gene product. *Development* **122**: 2673-85.

Bidwell, J.P., A.J. Van Wijnen, E.G. Fey, S. Dworetzky, S. Penman, J.L. Stein, J.B. Lian, and G.S. Stein. 1993. Osteocalcin gene promoter-binding factors are tissue-specific nuclear matrix components. *Proc Natl Acad Sci U S A* **90**: 3162-6.

Biggin, M.D. and W. McGinnis. 1997. Regulation of segmentation and segmental identity by Drosophila homeoproteins: the role of DNA binding in functional activity and specificity. *Development* **124**: 4425-33.

Black, B.E. and B.M. Paschal. 2004. Intranuclear organization and function of the androgen receptor. *Trends Endocrinol Metab* **15**: 411-7.

Black, B.E., M.J. Vitto, D. Gioeli, A. Spencer, N. Afshar, M.R. Conaway, M.J. Weber, and B.M. Paschal. 2004. Transient, ligand-dependent arrest of the androgen receptor in subnuclear foci alters phosphorylation and coactivator interactions. *Mol Endocrinol* **18**: 834-50.

Bober, E., T. Franz, H.H. Arnold, P. Gruss, and P. Tremblay. 1994. Pax-3 is required for the development of limb muscles: a possible role for the migration of dermomyotomal muscle progenitor cells. *Development* **120**: 603-12.

Bondurand, N., V. Pingault, D.E. Goerich, N. Lemort, E. Sock, C.L. Caignec, M. Wegner, and M. Goossens. 2000. Interaction among SOX10, PAX3 and MITF, three genes altered in Waardenburg syndrome. *Hum Mol Genet* **9**: 1907-17.

Bopp, D., M. Burri, S. Baumgartner, G. Frigerio, and M. Noll. 1986. Conservation of a large protein domain in the segmentation gene paired and in functionally related genes of Drosophila. *Cell* **47**: 1033-40.

Bosisio, D., I. Marazzi, A. Agresti, N. Shimizu, M.E. Bianchi, and G. Natoli. 2006. A hyper-dynamic equilibrium between promoter-bound and nucleoplasmic dimers controls NF-kappaB-dependent gene activity. *Embo J* **25**: 798-810.

Bottaro, D.P., J.S. Rubin, D.L. Faletto, A.M. Chan, T.E. Kmiecik, G.F. Vande Woude, and S.A. Aaronson. 1991. Identification of the hepatocyte growth factor receptor as the c-met proto-oncogene product. *Science* **251**: 802-4.

Bral, C.M., J.W. Steinke, C.J. Kang, and D.O. Peterson. 1998. RNA polymerase II transcription complex assembly in nuclear extracts. *Gene Expr* **7**: 191-204.

Bregman, D.B., L. Du, S. van der Zee, and S.L. Warren. 1995. Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J Cell Biol* **129**: 287-98.

Brickner, J.H. and P. Walter. 2004. Gene recruitment of the activated INO1 locus to the nuclear membrane. *PLoS Biol* **2**: e342.

Bridger, J.M., H. Herrmann, C. Munkel, and P. Lichter. 1998. Identification of an interchromosomal compartment by polymerization of nuclear-targeted vimentin. *J Cell Sci* **111 (Pt 9)**: 1241-53.

Brown, C.B., K.A. Engleka, J. Wenning, M. Min Lu, and J.A. Epstein. 2005. Identification of a hypaxial somite enhancer element regulating Pax3 expression in migrating myoblasts and characterization of hypaxial muscle Cre transgenic mice. *Genesis* **41**: 202-9.

Brown, K.E., S. Amoils, J.M. Horn, V.J. Buckle, D.R. Higgs, M. Merkenschlager, and A.G. Fisher. 2001. Expression of alpha- and beta-globin genes occurs within different nuclear domains in haemopoietic cells. *Nat Cell Biol* **3**: 602-6.

Bruun, J.A., E.I. Thomassen, K. Kristiansen, G. Tylden, T. Holm, I. Mikkola, G. Bjorkoy, and T. Johansen. 2005. The third helix of the homeodomain of paired class homeodomain proteins acts as a recognition helix both for DNA and protein interactions. *Nucleic Acids Res* **33**: 2661-75.

Buchberger, A., D. Freitag, and H.H. Arnold. 2007. A homeo-paired domain-binding motif directs Myf5 expression in progenitor cells of limb muscle. *Development* **134**: 1171-80.

Burri, M., Y. Tromvoukis, D. Bopp, G. Frigerio, and M. Noll. 1989. Conservation of the paired domain in metazoans and its structure in three isolated human genes. *Embo J* **8**: 1183-90.

Buscarlet, M. and S. Stifani. 2007. The 'Marx' of Groucho on development and disease. *Trends Cell Biol* **17**: 353-61.

Buttyan, R., C.A. Olsson, B. Sheard, and J. Kallos. 1983. Steroid receptor-nuclear matrix interactions. The role of DNA. *J Biol Chem* **258**: 14366-70.

Cai, Y., P.D. Brophy, I. Levitan, S. Stifani, and G.R. Dressler. 2003. Groucho suppresses Pax2 transactivation by inhibition of JNK-mediated phosphorylation. *Embo J* **22**: 5522-9.

Canon, J. and U. Banerjee. 2003. In vivo analysis of a developmental circuit for direct transcriptional activation and repression in the same cell by a Runx protein. *Genes Dev* **17**: 838-43.

Cao, Y. and C. Wang. 2000. The COOH-terminal transactivation domain plays a key role in regulating the in vitro and in vivo function of Pax3 homeodomain. *J Biol Chem* **275**: 9854-62.

Carroll, J.S., X.S. Liu, A.S. Brodsky, W. Li, C.A. Meyer, A.J. Szary, J. Eeckhoute, W. Shao, E.V. Hestermann, T.R. Geistlinger, E.A. Fox, P.A. Silver, and M. Brown. 2005. Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**: 33-43.

Casolari, J.M., C.R. Brown, D.A. Drubin, O.J. Rando, and P.A. Silver. 2005. Developmentally induced changes in transcriptional program alter spatial organization across chromosomes. *Genes Dev* **19**: 1188-98.

Casolari, J.M., C.R. Brown, S. Komili, J. West, H. Hieronymus, and P.A. Silver. 2004. Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* **117**: 427-39.

Chalepakis, G., R. Fritsch, H. Fickenscher, U. Deutsch, M. Goulding, and P. Gruss. 1991. The molecular basis of the undulated/Pax-1 mutation. *Cell* **66**: 873-84.

Chalepakis, G., M. Goulding, A. Read, T. Strachan, and P. Gruss. 1994a. Molecular basis of splotch and Waardenburg Pax-3 mutations. *Proc Natl Acad Sci U S A* **91**: 3685-9.

Chalepakis, G. and P. Gruss. 1995. Identification of DNA recognition sequences for the Pax3 paired domain. *Gene* **162**: 267-70.

Chalepakis, G., F.S. Jones, G.M. Edelman, and P. Gruss. 1994b. Pax-3 contains domains for transcription activation and transcription inhibition. *Proc Natl Acad Sci U S A* **91**: 12745-9.

Chalepakis, G., J. Wijnholds, and P. Gruss. 1994c. Pax-3-DNA interaction: flexibility in the DNA binding and induction of DNA conformational changes by paired domains. *Nucleic Acids Res* **22**: 3131-7.

Chambeyron, S. and W.A. Bickmore. 2004. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* **18**: 1119-30.

Chan, S.K. and R.S. Mann. 1993. The segment identity functions of Ultrabithorax are contained within its homeo domain and carboxy-terminal sequences. *Genes Dev* **7**: 796-811.

Chaney, B.A., K. Clark-Baldwin, V. Dave, J. Ma, and M. Rance. 2005. Solution structure of the K50 class homeodomain PITX2 bound to DNA and implications for mutations that cause Rieger syndrome. *Biochemistry* **44**: 7497-511.

Chang, C.P., Y. Jacobs, T. Nakamura, N.A. Jenkins, N.G. Copeland, and M.L. Cleary. 1997. Meis proteins are major in vivo DNA binding partners for wild type but not chimeric Pbx proteins. *Mol Cell Biol* **17**: 5679-87.

Chang, C.P., W.F. Shen, S. Rozenfeld, H.J. Lawrence, C. Largman, and M.L. Cleary. 1995. Pbx proteins display hexapeptide-dependent cooperative DNA binding with a subset of Hox proteins. *Genes Dev* **9**: 663-74.

Chen, D., M. Dundr, C. Wang, A. Leung, A. Lamond, T. Misteli, and S. Huang. 2005. Condensed mitotic chromatin is accessible to transcription factors and chromatin structural proteins. *J Cell Biol* **168**: 41-54.

Chen, D., C.S. Hinkley, R.W. Henry, and S. Huang. 2002. TBP dynamics in living human cells: constitutive association of TBP with mitotic chromosomes. *Mol Biol Cell* **13**: 276-84.

Chen, G. and A.J. Courey. 2000. Groucho/TLE family proteins and transcriptional repression. *Gene* **249**: 1-16.

Cheutin, T., A.J. McNairn, T. Jenuwein, D.M. Gilbert, P.B. Singh, and T. Misteli. 2003. Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* **299**: 721-5.

Chi, N. and J.A. Epstein. 2002. Getting your Pax straight: Pax proteins in development and disease. *Trends Genet* **18**: 41-7.

Chiang, C., Y. Litingtung, E. Lee, K.E. Young, J.L. Corden, H. Westphal, and P.A. Beachy. 1996. Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. *Nature* **383**: 407-13.

Choi, J.Y., J. Pratap, A. Javed, S.K. Zaidi, L. Xing, E. Balint, S. Dalamangas, B. Boyce, A.J. van Wijnen, J.B. Lian, J.L. Stein, S.N. Jones, and G.S. Stein. 2001. Subnuclear targeting of Runx/Cbfa/AML factors is essential for tissue-specific differentiation during embryonic development. *Proc Natl Acad Sci U S A* **98**: 8650-5.

Christova, R. and T. Oelgeschlager. 2002. Association of human TFIID-promoter complexes with silenced mitotic chromatin in vivo. *Nat Cell Biol* **4**: 79-82.

Chuang, C.H., A.E. Carpenter, B. Fuchsova, T. Johnson, P. de Lanerolle, and A.S. Belmont. 2006. Long-range directional movement of an interphase chromosome site. *Curr Biol* **16**: 825-31.

Chubb, J.R., S. Boyle, P. Perry, and W.A. Bickmore. 2002. Chromatin motion is constrained by association with nuclear compartments in human cells. *Curr Biol* **12**: 439-45.

Cirillo, L.A., F.R. Lin, I. Cuesta, D. Friedman, M. Jarnik, and K.S. Zaret. 2002. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell* **9**: 279-89.

Cmarko, D., P.J. Verschure, T.E. Martin, M.E. Dahmus, S. Krause, X.D. Fu, R. van Driel, and S. Fakan. 1999. Ultrastructural analysis of transcription and splicing in the cell nucleus after bromo-UTP microinjection. *Mol Biol Cell* **10**: 211-23.

Copeland, J.W., A. Nasiadka, B.H. Dietrich, and H.M. Krause. 1996. Patterning of the Drosophila embryo by a homeodomain-deleted Ftz polypeptide. *Nature* **379**: 162-5.

Copley, R.R. 2005. The EH1 motif in metazoan transcription factors. *BMC Genomics* **6**: 169.

Corry, G.N. and D.A. Underhill. 2005a. Pax3 target gene recognition occurs through distinct modes that are differentially affected by disease-associated mutations. *Pigment Cell Res* **18**: 427-38.

-. 2005b. Subnuclear compartmentalization of sequence-specific transcription factors and regulation of eukaryotic gene expression. *Biochem Cell Biol* **83**: 535-47.

Corsetti, M.T., G. Levi, F. Lancia, L. Sanseverino, S. Ferrini, E. Boncinelli, and G. Corte. 1995. Nucleolar localisation of three Hox homeoproteins. *J Cell Sci* **108 (Pt 1)**: 187-93.

Courey, A.J. and S. Jia. 2001. Transcriptional repression: the long and the short of it. *Genes Dev* **15**: 2786-96.

Cremer, T. and C. Cremer. 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* **2**: 292-301.

Cremer, T., M. Cremer, S. Dietzel, S. Muller, I. Solovei, and S. Fakan. 2006. Chromosome territories--a functional nuclear landscape. *Curr Opin Cell Biol* **18**: 307-16.

Cremer, T., G. Kreth, H. Koester, R.H. Fink, R. Heintzmann, M. Cremer, I. Solovei, D. Zink, and C. Cremer. 2000. Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. *Crit Rev Eukaryot Gene Expr* **10**: 179-212.

Cremer, T., K. Kupper, S. Dietzel, and S. Fakan. 2004. Higher order chromatin architecture in the cell nucleus: on the way from structure to function. *Biol Cell* **96**: 555-67.

Croft, J.A., J.M. Bridger, S. Boyle, P. Perry, P. Teague, and W.A. Bickmore. 1999. Differences in the localization and morphology of chromosomes in the human nucleus. *J Cell Biol* **145**: 1119-31.

Czerny, T., G. Schaffner, and M. Busslinger. 1993. DNA sequence recognition by Pax proteins: bipartite structure of the paired domain and its binding site. *Genes Dev* **7**: 2048-61.

Dahl, E., H. Koseki, and R. Balling. 1997. Pax genes and organogenesis. *Bioessays* **19**: 755-65.

Damante, G. and R. Di Lauro. 1991. Several regions of Antennapedia and thyroid transcription factor 1 homeodomains contribute to DNA binding specificity. *Proc Natl Acad Sci U S A* **88**: 5388-92.

da-Silva, E.O. 1991. Waardenburg I syndrome: a clinical and genetic study of two large Brazilian kindreds, and literature review. *Am J Med Genet* **40**: 65-74.

Daston, G., E. Lamar, M. Olivier, and M. Goulding. 1996. Pax-3 is necessary for migration but not differentiation of limb muscle precursors in the mouse. *Development* **122**: 1017-27.

Dave, V., C. Zhao, F. Yang, C.S. Tung, and J. Ma. 2000. Reprogrammable recognition codes in bicoid homeodomain-DNA interaction. *Mol Cell Biol* **20**: 7673-84.

Day, R.N., T.C. Voss, J.F. Enwright, 3rd, C.F. Booker, A. Periasamy, and F. Schaufele. 2003. Imaging the localized protein interactions between Pit-1 and the CCAAT/enhancer binding protein alpha in the living pituitary cell nucleus. *Mol Endocrinol* **17**: 333-45.

de Jong, L., M.A. Grande, K.A. Mattern, W. Schul, and R. van Driel. 1996. Nuclear domains involved in RNA synthesis, RNA processing, and replication. *Crit Rev Eukaryot Gene Expr* **6**: 215-46.

DeFranco, D.B. 1999. Regulation of steroid receptor subcellular trafficking. *Cell Biochem Biophys* **30**: 1-24.

D'Elia, A.V., G. Tell, I. Paron, L. Pellizzari, R. Lonigro, and G. Damante. 2001. Missense mutations of human homeoboxes: A review. *Hum Mutat* **18**: 361-74.

Denegri, M., D. Moralli, M. Rocchi, M. Biggiogera, E. Raimondi, F. Cobianchi, L. De Carli, S. Riva, and G. Biamonti. 2002. Human chromosomes 9, 12, and 15 contain the nucleation sites of stress-induced nuclear bodies. *Mol Biol Cell* **13**: 2069-79.

Dessain, S., C.T. Gross, M.A. Kuziora, and W. McGinnis. 1992. Antp-type homeodomains have distinct DNA binding specificities that correlate with their different regulatory functions in embryos. *Embo J* **11**: 991-1002.

DeStefano, A.L., L.A. Cupples, K.S. Arnos, J.H. Asher, Jr., C.T. Baldwin, S. Blanton, M.L. Carey, E.O. da Silva, T.B. Friedman, J. Greenberg, A.K. Lalwani, A. Milunsky, W.E. Nance, A. Pandya, R.S. Ramesar, A.P. Read, M. Tassabejhi, E.R. Wilcox, and L.A. Farrer. 1998. Correlation between Waardenburg syndrome phenotype and genotype in a population of individuals with identified PAX3 mutations. *Hum Genet* **102**: 499-506.

Downes, M., P. Ordentlich, H.Y. Kao, J.G. Alvarez, and R.M. Evans. 2000. Identification of a nuclear domain with deacetylase activity. *Proc Natl Acad Sci U S A* **97**: 10330-5.

Driever, W., J. Ma, C. Nusslein-Volhard, and M. Ptashne. 1989a. Rescue of bicoid mutant Drosophila embryos by bicoid fusion proteins containing heterologous activating sequences. *Nature* **342**: 149-54.

Driever, W., G. Thoma, and C. Nusslein-Volhard. 1989b. Determination of spatial domains of zygotic gene expression in the Drosophila embryo by the affinity of binding sites for the bicoid morphogen. *Nature* **340**: 363-7.

Dubnicoff, T., S.A. Valentine, G. Chen, T. Shi, J.A. Lengyel, Z. Paroush, and A.J. Courey. 1997. Conversion of dorsal from an activator to a repressor by the global corepressor Groucho. *Genes Dev* **11**: 2952-7.

Dundr, M. and T. Misteli. 2001. Functional architecture in the cell nucleus. *Biochem J* **356**: 297-310.

Dundr, M., J.K. Ospina, M.H. Sung, S. John, M. Upender, T. Ried, G.L. Hager, and A.G. Matera. 2007. Actin-dependent intranuclear repositioning of an active gene locus in vivo. *J Cell Biol* **179**: 1095-103.

Eberhard, D., G. Jimenez, B. Heavey, and M. Busslinger. 2000. Transcriptional repression by Pax5 (BSAP) through interaction with corepressors of the Groucho family. *Embo J* **19**: 2292-303.

Eberharter, A. and P.B. Becker. 2002. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO Rep* **3**: 224-9.

Eggenschwiler, J.T. and K.V. Anderson. 2000. Dorsal and lateral fates in the mouse neural tube require the cell-autonomous activity of the open brain gene. *Dev Biol* **227**: 648-60.

Eggenschwiler, J.T., E. Espinoza, and K.V. Anderson. 2001. Rab23 is an essential negative regulator of the mouse Sonic hedgehog signalling pathway. *Nature* **412**: 194-8.

Eilbracht, J. and M.S. Schmidt-Zachmann. 2001. Identification of a sequence element directing a protein to nuclear speckles. *Proc Natl Acad Sci U S A* **98**: 3849-54.

Eisen, T., D.J. Easty, D.C. Bennett, and C.R. Goding. 1995. The POU domain transcription factor Brn-2: elevated expression in malignant melanoma and regulation of melanocyte-specific gene expression. *Oncogene* **11**: 2157-64.

Ekker, S.C., D.P. von Kessler, and P.A. Beachy. 1992. Differential DNA sequence recognition is a determinant of specificity in homeotic gene action. *Embo J* **11**: 4059-72.

Elbi, C., D.A. Walker, G. Romero, W.P. Sullivan, D.O. Toft, G.L. Hager, and D.B. DeFranco. 2004. Molecular chaperones function as steroid receptor nuclear mobility factors. *Proc Natl Acad Sci U S A* **101**: 2876-81.

Elefanty, A.G., M. Antoniou, N. Custodio, M. Carmo-Fonseca, and F.G. Grosveld. 1996. GATA transcription factors associate with a novel class of nuclear bodies in erythroblasts and megakaryocytes. *Embo J* **15**: 319-33.

Engleka, K.A., A.D. Gitler, M. Zhang, D.D. Zhou, F.A. High, and J.A. Epstein. 2005. Insertion of Cre into the Pax3 locus creates a new allele of Splotch and identifies unexpected Pax3 derivatives. *Dev Biol* **280**: 396-406.

Englert, C., M. Vidal, S. Maheswaran, Y. Ge, R.M. Ezzell, K.J. Isselbacher, and D.A. Haber. 1995. Truncated WT1 mutants alter the subnuclear localization of the wild type protein. *Proc Natl Acad Sci U S A* **92**: 11960-4.

Enwright, J.F., 3rd, M.A. Kawecki-Crook, T.C. Voss, F. Schaufele, and R.N. Day. 2003. A PIT-1 homeodomain mutant blocks the intranuclear recruitment of the CCAAT/enhancer binding protein alpha required for prolactin gene transcription. *Mol Endocrinol* **17**: 209-22.

Epstein, D.J., D. Malo, M. Vekemans, and P. Gros. 1991a. Molecular characterization of a deletion encompassing the splotch mutation on mouse chromosome 1. *Genomics* **10**: 89-93.

Epstein, D.J., M. Vekemans, and P. Gros. 1991b. Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. *Cell* **67**: 767-74.

Epstein, D.J., K.J. Vogan, D.G. Trasler, and P. Gros. 1993. A mutation within intron 3 of the Pax-3 gene produces aberrantly spliced mRNA transcripts in the splotch (Sp) mouse mutant. *Proc Natl Acad Sci U S A* **90**: 532-6.

Epstein, J., J. Cai, T. Glaser, L. Jepeal, and R. Maas. 1994. Identification of a Pax paired domain recognition sequence and evidence for DNA-dependent conformational changes. *J Biol Chem* **269**: 8355-61.

Epstein, J.A., P. Lam, L. Jepeal, R.L. Maas, and D.N. Shapiro. 1995. Pax3 inhibits myogenic differentiation of cultured myoblast cells. *J Biol Chem* **270**: 11719-22.

Epstein, J.A., D.N. Shapiro, J. Cheng, P.Y. Lam, and R.L. Maas. 1996. Pax3 modulates expression of the c-Met receptor during limb muscle development. *Proc Natl Acad Sci U S A* **93**: 4213-8.

Epstein, J.A., B. Song, M. Lakkis, and C. Wang. 1998. Tumor-specific PAX3-FKHR transcription factor, but not PAX3, activates the platelet-derived growth factor alpha receptor. *Mol Cell Biol* **18**: 4118-30.

Evans, E.P., M.D. Burtenshaw, C.V. Beechey, and A.G. Searle. 1988. Mutations at the *Sp* locus. *Mouse News Letter* **75**: 28.

Fakan, S. 1994. Perichromatin fibrils are in situ forms of nascent transcripts. *Trends Cell Biol* **4**: 86-90.

-. 2004. The functional architecture of the nucleus as analysed by ultrastructural cytochemistry. *Histochem Cell Biol* **122**: 83-93.

Farla, P., R. Hersmus, B. Geverts, P.O. Mari, A.L. Nigg, H.J. Dubbink, J. Trapman, and A.B. Houtsmuller. 2004. The androgen receptor ligand-binding domain stabilizes DNA binding in living cells. *J Struct Biol* **147**: 50-61.

Farrer, L.A., K.S. Arnos, J.H. Asher, Jr., C.T. Baldwin, S.R. Diehl, T.B. Friedman, J. Greenberg, K.M. Grundfast, C. Hoth, A.K. Lalwani, and et al. 1994. Locus heterogeneity for Waardenburg syndrome is predictive of clinical subtypes. *Am J Hum Genet* **55**: 728-37.

Feister, H.A., K. Torrungruang, P. Thunyakitpisal, G.E. Parker, S.J. Rhodes, and J.P. Bidwell. 2000. NP/NMP4 transcription factors have distinct osteoblast nuclear matrix subdomains. *J Cell Biochem* **79**: 506-17.

Fejes-Toth, G., D. Pearce, and A. Naray-Fejes-Toth. 1998. Subcellular localization of mineralocorticoid receptors in living cells: effects of receptor agonists and antagonists. *Proc Natl Acad Sci U S A* **95**: 2973-8.

Festenstein, R., S.N. Pagakis, K. Hiragami, D. Lyon, A. Verreault, B. Sekkali, and D. Kioussis. 2003. Modulation of heterochromatin protein 1 dynamics in primary Mammalian cells. *Science* **299**: 719-21.

Fey, E.G., G. Krochmalnic, and S. Penman. 1986. The nonchromatin substructures of the nucleus: the ribonucleoprotein (RNP)-containing and RNP-depleted matrices analyzed by sequential fractionation and resinless section electron microscopy. *J Cell Biol* **102**: 1654-65.

Fisher, A.L. and M. Caudy. 1998. Groucho proteins: transcriptional corepressors for specific subsets of DNA-binding transcription factors in vertebrates and invertebrates. *Genes Dev* **12**: 1931-40.

Fitzpatrick, V.D., A. Percival-Smith, C.J. Ingles, and H.M. Krause. 1992. Homeodomain-independent activity of the fushi tarazu polypeptide in Drosophila embryos. *Nature* **356**: 610-2.

Fitzsimmons, D., W. Hodsdon, W. Wheat, S.M. Maira, B. Wasylyk, and J. Hagman. 1996. Pax-5 (BSAP) recruits Ets proto-oncogene family proteins to form functional ternary complexes on a B-cell-specific promoter. *Genes Dev* **10**: 2198-211.

Flores-Saaib, R.D., S. Jia, and A.J. Courey. 2001. Activation and repression by the C-terminal domain of Dorsal. *Development* **128**: 1869-79.

Fortin, A.S., D.A. Underhill, and P. Gros. 1997. Reciprocal effect of Waardenburg syndrome mutations on DNA binding by the Pax-3 paired domain and homeodomain. *Hum Mol Genet* **6**: 1781-90.

-. 1998. Helix 2 of the paired domain plays a key role in the regulation of DNA-binding by the Pax-3 homeodomain. *Nucleic Acids Res* **26**: 4574-81.

Francastel, C., W. Magis, and M. Groudine. 2001. Nuclear relocation of a transactivator subunit precedes target gene activation. *Proc Natl Acad Sci U S A* **98**: 12120-5.

Franz, T., R. Kothary, M.A. Surani, Z. Halata, and M. Grim. 1993. The Splotch mutation interferes with muscle development in the limbs. *Anat Embryol (Berl)* **187**: 153-60.

Frascella, E., L. Toffolatti, and A. Rosolen. 1998. Normal and rearranged PAX3 expression in human rhabdomyosarcoma. *Cancer Genet Cytogenet* **102**: 104-9.

Frazee, R.W., J.A. Taylor, and T.D. Tullius. 2002. Interchange of DNA-binding modes in the deformed and ultrabithorax homeodomains: a structural role for the N-terminal arm. *J Mol Biol* **323**: 665-83.

Fredericks, W.J., N. Galili, S. Mukhopadhyay, G. Rovera, J. Bennicelli, F.G. Barr, and F.J. Rauscher, 3rd. 1995. The PAX3-FKHR fusion protein created by the t(2;13) translocation in alveolar rhabdomyosarcomas is a more potent transcriptional activator than PAX3. *Mol Cell Biol* **15**: 1522-35.

Frigerio, G., M. Burri, D. Bopp, S. Baumgartner, and M. Noll. 1986. Structure of the segmentation gene paired and the Drosophila PRD gene set as part of a gene network. *Cell* **47**: 735-46.

Fujioka, M., P. Miskiewicz, L. Raj, A.A. Gulledge, M. Weir, and T. Goto. 1996. Drosophila Paired regulates late even-skipped expression through a composite binding site for the paired domain and the homeodomain. *Development* **122**: 2697-707.

Furukubo-Tokunaga, K., S. Flister, and W.J. Gehring. 1993. Functional specificity of the Antennapedia homeodomain. *Proc Natl Acad Sci U S A* **90**: 6360-4.

Furukubo-Tokunaga, K., M. Muller, M. Affolter, L. Pick, U. Kloter, and W.J. Gehring. 1992. In vivo analysis of the helix-turn-helix motif of the fushi tarazu homeo domain of Drosophila melanogaster. *Genes Dev* **6**: 1082-96.

Galibert, M.D., U. Yavuzer, T.J. Dexter, and C.R. Goding. 1999. Pax3 and regulation of the melanocyte-specific tyrosinase-related protein-1 promoter. *J Biol Chem* **274**: 26894-900.

Galili, N., R.J. Davis, W.J. Fredericks, S. Mukhopadhyay, F.J. Rauscher, 3rd, B.S. Emanuel, G. Rovera, and F.G. Barr. 1993. Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma. *Nat Genet* **5**: 230-5.

Garvie, C.W., J. Hagman, and C. Wolberger. 2001. Structural studies of Ets-1/Pax5 complex formation on DNA. *Mol Cell* **8**: 1267-76.

Gasperowicz, M. and F. Otto. 2005. Mammalian Groucho homologs: redundancy or specificity? *J Cell Biochem* **95**: 670-87.

Gehring, W.J., M. Affolter, and T. Burglin. 1994a. Homeodomain proteins. *Annu Rev Biochem* **63**: 487-526.

Gehring, W.J., Y.Q. Qian, M. Billeter, K. Furukubo-Tokunaga, A.F. Schier, D. Resendez-Perez, M. Affolter, G. Otting, and K. Wuthrich. 1994b. Homeodomain-DNA recognition. *Cell* **78**: 211-23.

Georget, V., J.M. Lobaccaro, B. Terouanne, P. Mangeat, J.C. Nicolas, and C. Sultan. 1997. Trafficking of the androgen receptor in living cells with fused green fluorescent protein-androgen receptor. *Mol Cell Endocrinol* **129**: 17-26.

Gibson, G., A. Schier, P. LeMotte, and W.J. Gehring. 1990. The specificities of Sex combs reduced and Antennapedia are defined by a distinct portion of each protein that includes the homeodomain. *Cell* **62**: 1087-103.

Giordani, J., L. Bajard, J. Demignon, P. Daubas, M. Buckingham, and P. Maire. 2007. Six proteins regulate the activation of Myf5 expression in embryonic mouse limbs. *Proc Natl Acad Sci U S A* **104**: 11310-5.

Glaser, T., D.S. Walton, and R.L. Maas. 1992. Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. *Nat Genet* **2**: 232-9.

Goding, C.R. 2000. Mitf from neural crest to melanoma: signal transduction and transcription in the melanocyte lineage. *Genes Dev* **14**: 1712-28.

Goulding, M.D., G. Chalepakis, U. Deutsch, J.R. Erselius, and P. Gruss. 1991. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *Embo J* **10**: 1135-47.

Goulding, M.D., A. Lumsden, and P. Gruss. 1993. Signals from the notochord and floor plate regulate the region-specific expression of two Pax genes in the developing spinal cord. *Development* **117**: 1001-16.

Grande, M.A., I. van der Kraan, L. de Jong, and R. van Driel. 1997. Nuclear distribution of transcription factors in relation to sites of transcription and RNA polymerase II. *J Cell Sci* **110 (Pt 15)**: 1781-91.

Grondin, B., M. Bazinet, and M. Aubry. 1996. The KRAB zinc finger gene ZNF74 encodes an RNA-binding protein tightly associated with the nuclear matrix. *J Biol Chem* **271**: 15458-67.

Groth, A., W. Rocha, A. Verreault, and G. Almouzni. 2007. Chromatin challenges during DNA replication and repair. *Cell* **128**: 721-33.

Grueneberg, D.A., S. Natesan, C. Alexandre, and M.Z. Gilman. 1992. Human and Drosophila homeodomain proteins that enhance the DNA-binding activity of serum response factor. *Science* **257**: 1089-95.

Grueneberg, D.A., K.J. Simon, K. Brennan, and M. Gilman. 1995. Sequence-specific targeting of nuclear signal transduction pathways by homeodomain proteins. *Mol Cell Biol* **15**: 3318-26.

Guo, B., P.R. Odgren, A.J. van Wijnen, T.J. Last, J. Nickerson, S. Penman, J.B. Lian, J.L. Stein, and G.S. Stein. 1995. The nuclear matrix protein NMP-1 is the transcription factor YY1. *Proc Natl Acad Sci U S A* **92**: 10526-30.

Hager, G.L., C. Elbi, and M. Becker. 2002. Protein dynamics in the nuclear compartment. *Curr Opin Genet Dev* **12**: 137-41.

Hager, G.L., C. Elbi, T.A. Johnson, T. Voss, A.K. Nagaich, R.L. Schiltz, Y. Qiu, and S. John. 2006. Chromatin dynamics and the evolution of alternate promoter states. *Chromosome Res* **14**: 107-16.

Hanes, S.D. and R. Brent. 1989. DNA specificity of the bicoid activator protein is determined by homeodomain recognition helix residue 9. *Cell* **57**: 1275-83.

Hanes, S.D., G. Riddihough, D. Ish-Horowicz, and R. Brent. 1994. Specific DNA recognition and intersite spacing are critical for action of the bicoid morphogen. *Mol Cell Biol* **14**: 3364-75.

Hanlon, S.E. and J.D. Lieb. 2004. Progress and challenges in profiling the dynamics of chromatin and transcription factor binding with DNA microarrays. *Curr Opin Genet Dev* **14**: 697-705.

Hansen, J.C. 2002. Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annu Rev Biophys Biomol Struct* **31**: 361-92.

Harris, R.G., E. White, E.S. Phillips, and K.A. Lillycrop. 2002. The expression of the developmentally regulated proto-oncogene Pax-3 is modulated by N-Myc. *J Biol Chem* **277**: 34815-25.

Haubst, N., J. Berger, V. Radjendirane, J. Graw, J. Favor, G.F. Saunders, A. Stoykova, and M. Gotz. 2004. Molecular dissection of Pax6 function: the specific roles of the paired domain and homeodomain in brain development. *Development* **131**: 6131-40.

Hayashi, S. and M.P. Scott. 1990. What determines the specificity of action of Drosophila homeodomain proteins? *Cell* **63**: 883-94.

He, S. and J.R. Davie. 2006. Sp1 and Sp3 foci distribution throughout mitosis. *J Cell Sci* **119**: 1063-70.

Heanue, T.A., R. Reshef, R.J. Davis, G. Mardon, G. Oliver, S. Tomarev, A.B. Lassar, and C.J. Tabin. 1999. Synergistic regulation of vertebrate muscle development by Dach2, Eya2, and Six1, homologs of genes required for Drosophila eye formation. *Genes Dev* **13**: 3231-43.

Hendzel, M.J., M.J. Kruhlak, N.A. MacLean, F. Boisvert, M.A. Lever, and D.P. Bazett-Jones. 2001. Compartmentalization of regulatory proteins in the cell nucleus. *J Steroid Biochem Mol Biol* **76**: 9-21.

Hewitt, S.L., F.A. High, S.L. Reiner, A.G. Fisher, and M. Merkenschlager. 2004. Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. *Eur J Immunol* **34**: 3604-13.

Hill, R.E., J. Favor, B.L. Hogan, C.C. Ton, G.F. Saunders, I.M. Hanson, J. Prosser, T. Jordan, N.D. Hastie, and V. van Heyningen. 1991. Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature* **354**: 522-5.

Hill, R.E., P.F. Jones, A.R. Rees, C.M. Sime, M.J. Justice, N.G. Copeland, N.A. Jenkins, E. Graham, and D.R. Davidson. 1989. A new family of mouse homeo box-containing genes: molecular structure, chromosomal location, and developmental expression of Hox-7.1. *Genes Dev* **3**: 26-37.

Hoey, T. and M. Levine. 1988. Divergent homeo box proteins recognize similar DNA sequences in Drosophila. *Nature* **332**: 858-61.

Hoey, T., R. Warrior, J. Manak, and M. Levine. 1988. DNA-binding activities of the Drosophila melanogaster even-skipped protein are mediated by its homeo domain and influenced by protein context. *Mol Cell Biol* **8**: 4598-607.

Hollenbach, A.D., C.J. McPherson, I. Lagutina, and G. Grosveld. 2002. The EF-hand calcium-binding protein calmyrin inhibits the transcriptional and DNA-binding activity of Pax3. *Biochim Biophys Acta* **1574**: 321-8.

Hollenbach, A.D., J.E. Sublett, C.J. McPherson, and G. Grosveld. 1999. The Pax3-FKHR oncoprotein is unresponsive to the Pax3-associated repressor hDaxx. *Embo J* **18**: 3702-11.

Hornyak, T.J., D.J. Hayes, L.Y. Chiu, and E.B. Ziff. 2001. Transcription factors in melanocyte development: distinct roles for Pax-3 and Mitf. *Mech Dev* **101**: 47-59.

Hoth, C.F., A. Milunsky, N. Lipsky, R. Sheffer, S.K. Clarren, and C.T. Baldwin. 1993. Mutations in the paired domain of the human PAX3 gene cause Klein-Waardenburg syndrome (WS-III) as well as Waardenburg syndrome type I (WS-I). *Am J Hum Genet* **52**: 455-62.

Hsieh, M.J., Y.L. Yao, I.L. Lai, and W.M. Yang. 2006. Transcriptional repression activity of PAX3 is modulated by competition between corepressor KAP1 and heterochromatin protein 1. *Biochem Biophys Res Commun* **349**: 573-81.

Htun, H., J. Barsony, I. Renyi, D.L. Gould, and G.L. Hager. 1996. Visualization of glucocorticoid receptor translocation and intranuclear organization in living cells with a green fluorescent protein chimera. *Proc Natl Acad Sci U S A* **93**: 4845-50.

Htun, H., L.T. Holth, D. Walker, J.R. Davie, and G.L. Hager. 1999. Direct visualization of the human estrogen receptor alpha reveals a role for ligand in the nuclear distribution of the receptor. *Mol Biol Cell* **10**: 471-86.

Iizuka, M. and M.M. Smith. 2003. Functional consequences of histone modifications. *Curr Opin Genet Dev* **13**: 154-60.

Isaac, V.E., P. Sciavolino, and C. Abate. 1995. Multiple amino acids determine the DNA binding specificity of the Msx-1 homeodomain. *Biochemistry* **34**: 7127-34.

Jackson, D.A. and P.R. Cook. 1985. Transcription occurs at a nucleoskeleton. *Embo J* **4**: 919-25.

Jackson, D.A., A.B. Hassan, R.J. Errington, and P.R. Cook. 1993. Visualization of focal sites of transcription within human nuclei. *Embo J* **12**: 1059-65.

Jackson, D.A., S.J. McCready, and P.R. Cook. 1981. RNA is synthesized at the nuclear cage. *Nature* **292**: 552-5.

Javed, A., B. Guo, S. Hiebert, J.Y. Choi, J. Green, S.C. Zhao, M.A. Osborne, S. Stifani, J.L. Stein, J.B. Lian, A.J. van Wijnen, and G.S. Stein. 2000. Groucho/TLE/R-esp proteins associate with the nuclear matrix and repress RUNX (CBF(alpha)/AML/PEBP2(alpha)) dependent activation of tissue-specific gene transcription. *J Cell Sci* **113 (Pt 12)**: 2221-31.

Jenuwein, T. and C.D. Allis. 2001. Translating the histone code. *Science* **293**: 1074-80.

Johnson, D.S., A. Mortazavi, R.M. Myers, and B. Wold. 2007. Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**: 1497-502.

Jolly, C., L. Konecny, D.L. Grady, Y.A. Kutskova, J.J. Cotto, R.I. Morimoto, and C. Vourc'h. 2002. In vivo binding of active heat shock transcription factor 1 to human chromosome 9 heterochromatin during stress. *J Cell Biol* **156**: 775-81.

Jun, S. and C. Desplan. 1996. Cooperative interactions between paired domain and homeodomain. *Development* **122**: 2639-50.

Kadonaga, J.T. 1990. Assembly and disassembly of the Drosophila RNA polymerase II complex during transcription. *J Biol Chem* **265**: 2624-31.

Kaffman, A. and E.K. O'Shea. 1999. Regulation of nuclear localization: a key to a door. *Annu Rev Cell Dev Biol* **15**: 291-339.

Kalionis, B. and P.H. O'Farrell. 1993. A universal target sequence is bound in vitro by diverse homeodomains. *Mech Dev* **43**: 57-70.

Kallunki, P., S. Jenkinson, G.M. Edelman, and F.S. Jones. 1995. Silencer elements modulate the expression of the gene for the neuron-glia cell adhesion molecule, Ng-CAM. *J Biol Chem* **270**: 21291-8.

Kamachi, Y., M. Uchikawa, A. Tanouchi, R. Sekido, and H. Kondoh. 2001. Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. *Genes Dev* **15**: 1272-86.

Kamakaka, R.T. and S. Biggins. 2005. Histone variants: deviants? *Genes Dev* **19**: 295-310.

Kanda, T., K.F. Sullivan, and G.M. Wahl. 1998. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr Biol* **8**: 377-85.

Karpova, T.S., T.Y. Chen, B.L. Sprague, and J.G. McNally. 2004. Dynamic interactions of a transcription factor with DNA are accelerated by a chromatin remodeller. *EMBO Rep* **5**: 1064-70.

Karvonen, U., O.A. Janne, and J.J. Palvimo. 2002. Pure antiandrogens disrupt the recruitment of coactivator GRIP1 to colocalize with androgen receptor in nuclei. *FEBS Lett* **523**: 43-7.

Kaufmann, S.H., S. Okret, A.C. Wikstrom, J.A. Gustafsson, and J.H. Shaper. 1986. Binding of the glucocorticoid receptor to the rat liver nuclear matrix. The role of disulfide bond formation. *J Biol Chem* **261**: 11962-7.

Kawate, H., Y. Wu, K. Ohnaka, R.H. Tao, K. Nakamura, T. Okabe, T. Yanase, H. Nawata, and R. Takayanagi. 2005. Impaired nuclear translocation, nuclear matrix targeting, and intranuclear mobility of mutant androgen receptors carrying amino acid substitutions in the deoxyribonucleic acid-binding domain derived from androgen insensitivity syndrome patients. *J Clin Endocrinol Metab* **90**: 6162-9.

Keller, C., B.R. Arenkiel, C.M. Coffin, N. El-Bardeesy, R.A. DePinho, and M.R. Capecchi. 2004a. Alveolar rhabdomyosarcomas in conditional Pax3:Fkhr mice: cooperativity of Ink4a/ARF and Trp53 loss of function. *Genes Dev* **18**: 2614-26.

Keller, C., M.S. Hansen, C.M. Coffin, and M.R. Capecchi. 2004b. Pax3:Fkhr interferes with embryonic Pax3 and Pax7 function: implications for alveolar rhabdomyosarcoma cell of origin. *Genes Dev* **18**: 2608-13.

Khan, J., M.L. Bittner, L.H. Saal, U. Teichmann, D.O. Azorsa, G.C. Gooden, W.J. Pavan, J.M. Trent, and P.S. Meltzer. 1999. cDNA microarrays detect activation of a myogenic transcription program by the PAX3-FKHR fusion oncogene. *Proc Natl Acad Sci U S A* **96**: 13264-9.

Kilchherr, F., S. Baumgartner, D. Bopp, E. Frei, and M. Noll. 1986. Isolation of the *paired* gene of *Drosophila* and its spatial expression during early embryogenesis. *Nature* **321**: 493-499.

Kimura, H., K. Sugaya, and P.R. Cook. 2002. The transcription cycle of RNA polymerase II in living cells. *J Cell Biol* **159**: 777-82.

Kino, T., S.H. Liou, E. Charmandari, and G.P. Chrousos. 2004. Glucocorticoid receptor mutants demonstrate increased motility inside the nucleus of living cells: time of fluorescence recovery after photobleaching (FRAP) is an integrated measure of receptor function. *Mol Med* **10**: 80-8.

Kioussi, C., M.K. Gross, and P. Gruss. 1995. Pax3: a paired domain gene as a regulator in PNS myelination. *Neuron* **15**: 553-62.

Kioussi, C. and P. Gruss. 1994. Differential induction of Pax genes by NGF and BDNF in cerebellar primary cultures. *J Cell Biol* **125**: 417-25.

Klemm, J.D., M.A. Rould, R. Aurora, W. Herr, and C.O. Pabo. 1994. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell* **77**: 21-32.

Koop, K.E., L.M. MacDonald, and C.G. Lobe. 1996. Transcripts of Grg4, a murine groucho-related gene, are detected in adjacent tissues to other murine neurogenic gene homologues during embryonic development. *Mech Dev* **59**: 73-87.

Kosak, S.T., J.A. Skok, K.L. Medina, R. Riblet, M.M. Le Beau, A.G. Fisher, and H. Singh. 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* **296**: 158-62.

Kozmik, Z., S. Wang, P. Dorfler, B. Adams, and M. Busslinger. 1992. The promoter of the CD19 gene is a target for the B-cell-specific transcription factor BSAP. *Mol Cell Biol* **12**: 2662-72.

Krogan, N.J., M. Kim, A. Tong, A. Golshani, G. Cagney, V. Canadien, D.P. Richards, B.K. Beattie, A. Emili, C. Boone, A. Shilatifard, S. Buratowski, and J. Greenblatt. 2003. Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* **23**: 4207-18.

Kuang, S., S.B. Charge, P. Seale, M. Huh, and M.A. Rudnicki. 2006. Distinct roles for Pax7 and Pax3 in adult regenerative myogenesis. *J Cell Biol* **172**: 103-13.

Kuhlbrodt, K., B. Herbarth, E. Sock, I. Hermans-Borgmeyer, and M. Wegner. 1998. Sox10, a novel transcriptional modulator in glial cells. *J Neurosci* **18**: 237-50.

Kumar, R., R. Betney, J. Li, E.B. Thompson, and I.J. McEwan. 2004. Induced alpha-helix structure in AF1 of the androgen receptor upon binding transcription factor TFIIF. *Biochemistry* **43**: 3008-13.

Kumar, R. and E.B. Thompson. 2003. Transactivation functions of the N-terminal domains of nuclear hormone receptors: protein folding and coactivator interactions. *Mol Endocrinol* **17**: 1-10.

Kurz, A., S. Lampel, J.E. Nickolenko, J. Bradl, A. Benner, R.M. Zirbel, T. Cremer, and P. Lichter. 1996. Active and inactive genes localize preferentially in the periphery of chromosome territories. *J Cell Biol* **135**: 1195-205.

Kuziora, M.A. and W. McGinnis. 1989. A homeodomain substitution changes the regulatory specificity of the deformed protein in Drosophila embryos. *Cell* **59**: 563-71.

Kwang, S.J., S.M. Brugger, A. Lazik, A.E. Merrill, L.Y. Wu, Y.H. Liu, M. Ishii, F.O. Sangiorgi, M. Rauchman, H.M. Sucov, R.L. Maas, and R.E. Maxson, Jr. 2002. Msx2 is an immediate downstream effector of Pax3 in the development of the murine cardiac neural crest. *Development* **129**: 527-38.

Lachner, M., D. O'Carroll, S. Rea, K. Mechtler, and T. Jenuwein. 2001. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**: 116-20.

Lagutina, I., S.J. Conway, J. Sublett, and G.C. Grosveld. 2002. Pax3-FKHR knock-in mice show developmental aberrations but do not develop tumors. *Mol Cell Biol* **22**: 7204-16.

Lamond, A.I. and W.C. Earnshaw. 1998. Structure and function in the nucleus. *Science* **280**: 547-53.

Lan, Y., M. Fujioka, R. Polsgrove, P. Miskiewicz, D. Morrissey, T. Goto, and M. Weir. 1998. Plasticity of Drosophila paired function. *Dev Genet* **23**: 45-55.

Lanctot, C., T. Cheutin, M. Cremer, G. Cavalli, and T. Cremer. 2007. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nat Rev Genet* **8**: 104-15.

Lang, D., F. Chen, R. Milewski, J. Li, M.M. Lu, and J.A. Epstein. 2000. Pax3 is required for enteric ganglia formation and functions with Sox10 to modulate expression of c-ret. *J Clin Invest* **106**: 963-71.

Lang, D. and J.A. Epstein. 2003. Sox10 and Pax3 physically interact to mediate activation of a conserved c-RET enhancer. *Hum Mol Genet* **12**: 937-45.

Lang, D., M.M. Lu, L. Huang, K.A. Engleka, M. Zhang, E.Y. Chu, S. Lipner, A. Skoultchi, S.E. Millar, and J.A. Epstein. 2005. Pax3 functions at a nodal point in melanocyte stem cell differentiation. *Nature* **433**: 884-7.

Lange, A., R.E. Mills, C.J. Lange, M. Stewart, S.E. Devine, and A.H. Corbett. 2007. Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem* **282**: 5101-5.

Larsson, S.H., J.P. Charlieu, K. Miyagawa, D. Engelkamp, M. Rassoulzadegan, A. Ross, F. Cuzin, V. van Heyningen, and N.D. Hastie. 1995. Subnuclear localization of WT1 in splicing or transcription factor domains is regulated by alternative splicing. *Cell* **81**: 391-401.

Lee, H., R. Habas, and C. Abate-Shen. 2004. MSX1 cooperates with histone H1b for inhibition of transcription and myogenesis. *Science* **304**: 1675-8.

Lee, H., J.C. Quinn, K.V. Prasanth, V.A. Swiss, K.D. Economides, M.M. Camacho, D.L. Spector, and C. Abate-Shen. 2006. PIAS1 confers DNA-binding specificity on the Msx1 homeoprotein. *Genes Dev* **20**: 784-94.

Lee, M., J. Goodall, C. Verastegui, R. Ballotti, and C.R. Goding. 2000. Direct regulation of the Microphthalmia promoter by Sox10 links Waardenburg-Shah syndrome (WS4)-associated hypopigmentation and deafness to WS2. *J Biol Chem* **275**: 37978-83.

Lee, T.I. and R.A. Young. 2000. Transcription of eukaryotic protein-coding genes. *Annu Rev Genet* **34**: 77-137.

Lefstin, J.A. and K.R. Yamamoto. 1998. Allosteric effects of DNA on transcriptional regulators. *Nature* **392**: 885-8.

Lemon, B. and R. Tjian. 2000. Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev* **14**: 2551-69.

Lever, M.A., J.P. Th'ng, X. Sun, and M.J. Hendzel. 2000. Rapid exchange of histone H1.1 on chromatin in living human cells. *Nature* **408**: 873-6.

Levi, V., Q. Ruan, M. Plutz, A.S. Belmont, and E. Gratton. 2005. Chromatin dynamics in interphase cells revealed by tracking in a two-photon excitation microscope. *Biophys J* **89**: 4275-85.

Li, B., M. Carey, and J.L. Workman. 2007. The role of chromatin during transcription. *Cell* **128**: 707-19.

Li, B., L. Howe, S. Anderson, J.R. Yates, 3rd, and J.L. Workman. 2003. The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **278**: 8897-903.

Li, G., G. Sudlow, and A.S. Belmont. 1998. Interphase cell cycle dynamics of a late-replicating, heterochromatic homogeneously staining region: precise choreography of condensation/decondensation and nuclear positioning. *J Cell Biol* **140**: 975-89.

Li, J., F. Chen, and J.A. Epstein. 2000. Neural crest expression of Cre recombinase directed by the proximal Pax3 promoter in transgenic mice. *Genesis* **26**: 162-4.

Li, J., K.C. Liu, F. Jin, M.M. Lu, and J.A. Epstein. 1999a. Transgenic rescue of congenital heart disease and spina bifida in Splotch mice. *Development* **126**: 2495-503.

Li, Q., U. Bjork, and O. Wrange. 1999b. Assays for interaction of transcription factor with nucleosome. *Methods Enzymol* **304**: 313-32.

Liem, K.F., Jr., G. Tremml, H. Roelink, and T.M. Jessell. 1995. Dorsal differentiation of neural plate cells induced by BMP-mediated signals from epidermal ectoderm. *Cell* **82**: 969-79.

Lin, L. and W. McGinnis. 1992. Mapping functional specificity in the Dfd and Ubx homeo domains. *Genes Dev* **6**: 1071-81.

Lira, S.A., E.B. Crenshaw, 3rd, C.K. Glass, L.W. Swanson, and M.G. Rosenfeld. 1988. Identification of rat growth hormone genomic sequences targeting pituitary expression in transgenic mice. *Proc Natl Acad Sci U S A* **85**: 4755-9.

Lohrum, M.A., M. Ashcroft, M.H. Kubbutat, and K.H. Vousden. 2000. Identification of a cryptic nucleolar-localization signal in MDM2. *Nat Cell Biol* **2**: 179-81.

Lorain, S., J.P. Quivy, F. Monier-Gavelle, C. Scamps, Y. Lecluse, G. Almouzni, and M. Lipinski. 1998. Core histones and HIRIP3, a novel histone-binding protein, directly interact with WD repeat protein HIRA. *Mol Cell Biol* **18**: 5546-56.

Lowings, P., U. Yavuzer, and C.R. Goding. 1992. Positive and negative elements regulate a melanocyte-specific promoter. *Mol Cell Biol* **12**: 3653-62.

Luger, K., A.W. Mader, R.K. Richmond, D.F. Sargent, and T.J. Richmond. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251-60.

Lukacs, G.L., P. Haggie, O. Seksek, D. Lechardeur, N. Freedman, and A.S. Verkman. 2000. Size-dependent DNA mobility in cytoplasm and nucleus. *J Biol Chem* **275**: 1625-9.

Lundgren, M., C.M. Chow, P. Sabbattini, A. Georgiou, S. Minaee, and N. Dillon. 2000. Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. *Cell* **103**: 733-43.

Ma, H., A.J. Siegel, and R. Berezney. 1999. Association of chromosome territories with the nuclear matrix. Disruption of human chromosome territories correlates with the release of a subset of nuclear matrix proteins. *J Cell Biol* **146**: 531-42.

Macina, R.A., F.G. Barr, N. Galili, and H.C. Riethman. 1995. Genomic organization of the human PAX3 gene: DNA sequence analysis of the region disrupted in alveolar rhabdomyosarcoma. *Genomics* **26**: 1-8.

Magnaghi, P., C. Roberts, S. Lorain, M. Lipinski, and P.J. Scambler. 1998. HIRA, a mammalian homologue of Saccharomyces cerevisiae transcriptional co-repressors, interacts with Pax3. *Nat Genet* **20**: 74-7.

Mahy, N.L., P.E. Perry, and W.A. Bickmore. 2002a. Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *J Cell Biol* **159**: 753-63.

Mahy, N.L., P.E. Perry, S. Gilchrist, R.A. Baldock, and W.A. Bickmore. 2002b. Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. *J Cell Biol* **157**: 579-89.

Mancini, M.A., B. Shan, J.A. Nickerson, S. Penman, and W.H. Lee. 1994. The retinoblastoma gene product is a cell cycle-dependent, nuclear matrix-associated protein. *Proc Natl Acad Sci U S A* **91**: 418-22.

Mancini, M.G., B. Liu, Z.D. Sharp, and M.A. Mancini. 1999. Subnuclear partitioning and functional regulation of the Pit-1 transcription factor. *J Cell Biochem* **72**: 322-38.

Mankoo, B.S., N.S. Collins, P. Ashby, E. Grigorieva, L.H. Pevny, A. Candia, C.V. Wright, P.W. Rigby, and V. Pachnis. 1999. Mox2 is a component of the genetic hierarchy controlling limb muscle development. *Nature* **400**: 69-73.

Mansouri, A., M. Hallonet, and P. Gruss. 1996. Pax genes and their roles in cell differentiation and development. *Curr Opin Cell Biol* **8**: 851-7.

Marmorstein, R. 2001. Protein modules that manipulate histone tails for chromatin regulation. *Nat Rev Mol Cell Biol* **2**: 422-32.

Maroto, M., R. Reshef, A.E. Munsterberg, S. Koester, M. Goulding, and A.B. Lassar. 1997. Ectopic Pax-3 activates MyoD and Myf-5 expression in embryonic mesoderm and neural tissue. *Cell* **89**: 139-48.

Marshall, W.F., A. Straight, J.F. Marko, J. Swedlow, A. Dernburg, A. Belmont, A.W. Murray, D.A. Agard, and J.W. Sedat. 1997. Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr Biol* **7**: 930-9.

Martelli, A.M., E. Falcieri, M. Zweyer, R. Bortul, G. Tabellini, A. Cappellini, L. Cocco, and L. Manzoli. 2002. The controversial nuclear matrix: a balanced point of view. *Histol Histopathol* **17**: 1193-205.

Martin, S. and A. Pombo. 2003. Transcription factories: quantitative studies of nanostructures in the mammalian nucleus. *Chromosome Res* **11**: 461-70.

Mayanil, C.S., D. George, L. Freilich, E.J. Miljan, B. Mania-Farnell, D.G. McLone, and E.G. Bremer. 2001. Microarray analysis detects novel Pax3 downstream target genes. *J Biol Chem* **276**: 49299-309.

Mayanil, C.S., A. Pool, H. Nakazaki, A.C. Reddy, B. Mania-Farnell, B. Yun, D. George, D.G. McLone, and E.G. Bremer. 2006. Regulation of murine TGFbeta2 by Pax3 during early embryonic development. *J Biol Chem* **281**: 24544-52.

McKenna, N.J., R.B. Lanz, and B.W. O'Malley. 1999. Nuclear receptor coregulators: cellular and molecular biology. *Endocr Rev* **20**: 321-44.

McKinnell, I.W., J. Ishibashi, F. Le Grand, V.G. Punch, G.C. Addicks, J.F. Greenblatt, F.J. Dilworth, and M.A. Rudnicki. 2008. Pax7 activates myogenic genes by recruitment of a histone methyltransferase complex. *Nat Cell Biol* **10**: 77-84.

McNally, J.G., W.G. Muller, D. Walker, R. Wolford, and G.L. Hager. 2000. The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* **287**: 1262-5.

McNeil, S., B. Guo, J.L. Stein, J.B. Lian, S. Bushmeyer, E. Seto, M.L. Atchison, S. Penman, A.J. van Wijnen, and G.S. Stein. 1998. Targeting of the YY1 transcription factor to the nucleolus and the nuclear matrix in situ: the C-terminus is a principal determinant for nuclear trafficking. *J Cell Biochem* **68**: 500-10.

Mennerich, D. and T. Braun. 2001. Activation of myogenesis by the homeobox gene Lbx1 requires cell proliferation. *Embo J* **20**: 7174-83.

Merriman, H.L., A.J. van Wijnen, S. Hiebert, J.P. Bidwell, E. Fey, J. Lian, J. Stein, and G.S. Stein. 1995. The tissue-specific nuclear matrix protein, NMP-2, is a member of the AML/CBF/PEBP2/runt domain transcription factor family: interactions with the osteocalcin gene promoter. *Biochemistry* **34**: 13125-32.

Meshorer, E. and T. Misteli. 2006. Chromatin in pluripotent embryonic stem cells and differentiation. *Nat Rev Mol Cell Biol* **7**: 540-6.

Mikkola, I., J.A. Bruun, T. Holm, and T. Johansen. 2001. Superactivation of Pax6-mediated transactivation from paired domain-binding sites by dna-independent recruitment of different homeodomain proteins. *J Biol Chem* **276**: 4109-18.

Milewski, R.C., N.C. Chi, J. Li, C. Brown, M.M. Lu, and J.A. Epstein. 2004. Identification of minimal enhancer elements sufficient for Pax3 expression in neural crest and implication of Tead2 as a regulator of Pax3. *Development* **131**: 829-37.

Milili, M., L. Gauthier, J. Veran, M.G. Mattei, and C. Schiff. 2002. A new Groucho TLE4 protein may regulate the repressive activity of Pax5 in human B lymphocytes. *Immunology* **106**: 447-55.

Mishra, R., I.P. Gorlov, L.Y. Chao, S. Singh, and G.F. Saunders. 2002. PAX6, paired domain influences sequence recognition by the homeodomain. *J Biol Chem* **277**: 49488-94.

Miskiewicz, P., D. Morrissey, Y. Lan, L. Raj, S. Kessler, M. Fujioka, T. Goto, and M. Weir. 1996. Both the paired domain and homeodomain are required for in vivo function of Drosophila Paired. *Development* **122**: 2709-18.

Misteli, T. 2001. Protein dynamics: implications for nuclear architecture and gene expression. *Science* **291**: 843-7.

-. 2007. Beyond the sequence: cellular organization of genome function. *Cell* **128**: 787-800.

Misteli, T., A. Gunjan, R. Hock, M. Bustin, and D.T. Brown. 2000. Dynamic binding of histone H1 to chromatin in living cells. *Nature* **408**: 877-81.

Moen, P.T., Jr., C.V. Johnson, M. Byron, L.S. Shopland, I.L. de la Serna, A.N. Imbalzano, and J.B. Lawrence. 2004. Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis. *Mol Biol Cell* **15**: 197-206.

Moorefield, K.S., H. Yin, T.D. Nichols, C. Cathcart, S.O. Simmons, and J.M. Horowitz. 2006. Sp2 localizes to subnuclear foci associated with the nuclear matrix. *Mol Biol Cell* **17**: 1711-22.

Morell, R., T.B. Friedman, S. Moeljopawiro, Hartono, Soewito, and J.H. Asher, Jr. 1992. A frameshift mutation in the HuP2 paired domain of the probable human homolog of murine Pax-3 is responsible for Waardenburg syndrome type 1 in an Indonesian family. *Hum Mol Genet* **1**: 243-7.

Muller, W.G., D. Walker, G.L. Hager, and J.G. McNally. 2001. Large-scale chromatin decondensation and recondensation regulated by transcription from a natural promoter. *J Cell Biol* **154**: 33-48.

Nagaich, A.K., D.A. Walker, R. Wolford, and G.L. Hager. 2004. Rapid periodic binding and displacement of the glucocorticoid receptor during chromatin remodeling. *Mol Cell* **14**: 163-74.

Natoli, T.A., M.K. Ellsworth, C. Wu, K.W. Gross, and S.C. Pruitt. 1997. Positive and negative DNA sequence elements are required to establish the pattern of Pax3 expression. *Development* **124**: 617-26.

Nazareth, L.V., D.L. Stenoien, W.E. Bingman, 3rd, A.J. James, C. Wu, Y. Zhang, D.P. Edwards, M. Mancini, M. Marcelli, D.J. Lamb, and N.L. Weigel. 1999. A C619Y mutation in the human androgen receptor causes inactivation and mislocalization of the receptor with concomitant sequestration of SRC-1 (steroid receptor coactivator 1). *Mol Endocrinol* **13**: 2065-75.

Ng, H.H., F. Robert, R.A. Young, and K. Struhl. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11**: 709-19.

Nickerson, J. 2001. Experimental observations of a nuclear matrix. *J Cell Sci* **114**: 463-74.

Nickerson, J.A., G. Krochmalnic, K.M. Wan, and S. Penman. 1989. Chromatin architecture and nuclear RNA. *Proc Natl Acad Sci U S A* **86**: 177-81.

Nusslein-Volhard, C. and E. Wieschaus. 1980. Mutations affecting segment number and polarity in Drosophila. *Nature* **287**: 795-801.

Nutt, S.L. and M. Busslinger. 1999. Monoallelic expression of Pax5: a paradigm for the haploinsufficiency of mammalian Pax genes? *Biol Chem* **380**: 601-11.

Okladnova, O., Y.V. Syagailo, M. Tranitz, P. Riederer, G. Stober, R. Mossner, and K.P. Lesch. 1999. Functional characterization of the human PAX3 gene regulatory region. *Genomics* **57**: 110-9.

Olson, M.O., M. Dundr, and A. Szebeni. 2000. The nucleolus: an old factory with unexpected capabilities. *Trends Cell Biol* **10**: 189-96.

Osborne, C.S., L. Chakalova, K.E. Brown, D. Carter, A. Horton, E. Debrand, B. Goyenechea, J.A. Mitchell, S. Lopes, W. Reik, and P. Fraser. 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* **36**: 1065-71.

O'Toole, P.J., T. Inoue, L. Emerson, I.E. Morrison, A.R. Mackie, R.J. Cherry, and J.D. Norton. 2003. Id proteins negatively regulate basic helix-loop-helix transcription factor function by disrupting subnuclear compartmentalization. *J Biol Chem* **278**: 45770-6.

Pachnis, V., B. Mankoo, and F. Costantini. 1993. Expression of the c-ret proto-oncogene during mouse embryogenesis. *Development* **119**: 1005-17.

Parker, C.J., S.G. Shawcross, H. Li, Q.Y. Wang, C.S. Herrington, S. Kumar, R.M. MacKie, W. Prime, I.G. Rennie, K. Sisley, and P. Kumar. 2004. Expression of PAX 3 alternatively spliced transcripts and identification of two new isoforms in human tumors of neural crest origin. *Int J Cancer* **108**: 314-20.

Parker, G.E., R.M. Sandoval, H.A. Feister, J.P. Bidwell, and S.J. Rhodes. 2000. The homeodomain coordinates nuclear entry of the Lhx3 neuroendocrine transcription factor and association with the nuclear matrix. *J Biol Chem* **275**: 23891-8.

Paroush, Z., R.L. Finley, Jr., T. Kidd, S.M. Wainwright, P.W. Ingham, R. Brent, and D. Ish-Horowicz. 1994. Groucho is required for Drosophila neurogenesis, segmentation, and sex determination and interacts directly with hairy-related bHLH proteins. *Cell* **79**: 805-15.

Pearce, D., A. Naray-Fejes-Toth, and G. Fejes-Toth. 2002. Determinants of subnuclear organization of mineralocorticoid receptor characterized through analysis of wild type and mutant receptors. *J Biol Chem* **277**: 1451-6.

Peters, A.H., J.E. Mermoud, D. O'Carroll, M. Pagani, D. Schweizer, N. Brockdorff, and T. Jenuwein. 2002. Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. *Nat Genet* **30**: 77-80.

Peterson, C.L. and M.A. Laniel. 2004. Histones and histone modifications. *Curr Biol* **14**: R546-51.

Pevny, L.H. and R. Lovell-Badge. 1997. Sox genes find their feet. *Curr Opin Genet Dev* **7**: 338-44.

Phair, R.D. and T. Misteli. 2000. High mobility of proteins in the mammalian cell nucleus. *Nature* **404**: 604-9.

Phair, R.D., P. Scaffidi, C. Elbi, J. Vecerova, A. Dey, K. Ozato, D.T. Brown, G. Hager, M. Bustin, and T. Misteli. 2004. Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. *Mol Cell Biol* **24**: 6393-402.

Pombo, A., P. Cuello, W. Schul, J.B. Yoon, R.G. Roeder, P.R. Cook, and S. Murphy. 1998. Regional and temporal specialization in the nucleus: a transcriptionally-active nuclear domain rich in PTF, Oct1 and PIKA antigens associates with specific chromosomes early in the cell cycle. *Embo J* **17**: 1768-78.

Pombo, A., E. Jones, F.J. Iborra, H. Kimura, K. Sugaya, P.R. Cook, and D.A. Jackson. 2000. Specialized transcription factories within mammalian nuclei. *Crit Rev Eukaryot Gene Expr* **10**: 21-9.

Pomerantz, J.L. and P.A. Sharp. 1994. Homeodomain determinants of major groove recognition. *Biochemistry* **33**: 10851-8.

Potterf, S.B., M. Furumura, K.J. Dunn, H. Arnheiter, and W.J. Pavan. 2000. Transcription factor hierarchy in Waardenburg syndrome: regulation of MITF expression by SOX10 and PAX3. *Hum Genet* **107**: 1-6.

Price, R.H., Jr., C.A. Butler, P. Webb, R. Uht, P. Kushner, and R.J. Handa. 2001. A splice variant of estrogen receptor beta missing exon 3 displays altered subnuclear localization and capacity for transcriptional activation. *Endocrinology* **142**: 2039-49.

Priston, M., K. Kozlowski, D. Gill, K. Letwin, Y. Buys, A.V. Levin, M.A. Walter, and E. Heon. 2001. Functional analyses of two newly identified PITX2 mutants reveal a novel molecular mechanism for Axenfeld-Rieger syndrome. *Hum Mol Genet* 10: 1631-8.

Pufall, M.A., G.M. Lee, M.L. Nelson, H.S. Kang, A. Velyvis, L.E. Kay, L.P. McIntosh, and B.J. Graves. 2005. Variable control of Ets-1 DNA binding by multiple phosphates in an unstructured region. *Science* **309**: 142-5.

Punzo, C., S. Kurata, and W.J. Gehring. 2001. The eyeless homeodomain is dispensable for eye development in Drosophila. *Genes Dev* **15**: 1716-23.

Ragoczy, T., M.A. Bender, A. Telling, R. Byron, and M. Groudine. 2006. The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. *Genes Dev* **20**: 1447-57.

Rajaram, N. and T.K. Kerppola. 2004. Synergistic transcription activation by Maf and Sox and their subnuclear localization are disrupted by a mutation in Maf that causes cataract. *Mol Cell Biol* **24**: 5694-709.

Ray, R.P., K. Arora, C. Nusslein-Volhard, and W.M. Gelbart. 1991. The control of cell fate along the dorsal-ventral axis of the Drosophila embryo. *Development* **113**: 35-54.

Read, A.P. and V.E. Newton. 1997. Waardenburg syndrome. *J Med Genet* **34**: 656-65.

Reichenzeller, M., A. Burzlaff, P. Lichter, and H. Herrmann. 2000. In vivo observation of a nuclear channel-like system: evidence for a distinct interchromosomal domain compartment in interphase cells. *J Struct Biol* **129**: 175-85.

Relaix, F., D. Montarras, S. Zaffran, B. Gayraud-Morel, D. Rocancourt, S. Tajbakhsh, A. Mansouri, A. Cumano, and M. Buckingham. 2006. Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J Cell Biol* **172**: 91-102.

Relaix, F., M. Polimeni, D. Rocancourt, C. Ponzetto, B.W. Schafer, and M. Buckingham. 2003. The transcriptional activator PAX3-FKHR rescues the defects of Pax3 mutant mice but induces a myogenic gain-of-function phenotype with ligand-independent activation of Met signaling in vivo. *Genes Dev* **17**: 2950-65.

Relaix, F., D. Rocancourt, A. Mansouri, and M. Buckingham. 2004. Divergent functions of murine Pax3 and Pax7 in limb muscle development. *Genes Dev* **18**: 1088-105.

Remenyi, A., A. Tomilin, E. Pohl, K. Lins, A. Philippsen, R. Reinbold, H.R. Scholer, and M. Wilmanns. 2001. Differential dimer activities of the transcription factor Oct-1 by DNA-induced interface swapping. *Mol Cell* **8**: 569-80.

Reshef, R., M. Maroto, and A.B. Lassar. 1998. Regulation of dorsal somitic cell fates: BMPs and Noggin control the timing and pattern of myogenic regulator expression. *Genes Dev* **12**: 290-303.

Riz, I., S.S. Akimov, S.S. Eaker, K.K. Baxter, H.J. Lee, L. Marino-Ramirez, D. Landsman, T.S. Hawley, and R.G. Hawley. 2007. TLX1/HOX11-induced hematopoietic differentiation blockade. *Oncogene* **26**: 4115-23.

Robinett, C.C., A. Straight, G. Li, C. Willhelm, G. Sudlow, A. Murray, and A.S. Belmont. 1996. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol* **135**: 1685-700.

Robson, E.J., S.J. He, and M.R. Eccles. 2006. A PANorama of PAX genes in cancer and development. *Nat Rev Cancer* **6**: 52-62.

Romeo, G., P. Ronchetto, Y. Luo, V. Barone, M. Seri, I. Ceccherini, B. Pasini, R. Bocciardi, M. Lerone, H. Kaariainen, and et al. 1994. Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. *Nature* **367**: 377-8.

Rubbi, C.P. and J. Milner. 2000. Non-activated p53 co-localizes with sites of transcription within both the nucleoplasm and the nucleolus. *Oncogene* **19**: 85-96.

Sanford, L.P., I. Ormsby, A.C. Gittenberger-de Groot, H. Sariola, R. Friedman, G.P. Boivin, E.L. Cardell, and T. Doetschman. 1997. TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes. *Development* **124**: 2659-70.

Santos-Rosa, H., R. Schneider, A.J. Bannister, J. Sherriff, B.E. Bernstein, N.C. Emre, S.L. Schreiber, J. Mellor, and T. Kouzarides. 2002. Active genes are tri-methylated at K4 of histone H3. *Nature* **419**: 407-11.

Sarge, K.D., S.P. Murphy, and R.I. Morimoto. 1993. Activation of heat shock gene transcription by heat shock factor 1 involves oligomerization, acquisition of DNA-binding activity, and nuclear localization and can occur in the absence of stress. *Mol Cell Biol* **13**: 1392-407.

Sarge, K.D. and O.K. Park-Sarge. 2005. Gene bookmarking: keeping the pages open. *Trends Biochem Sci* **30**: 605-10.

Sarma, K. and D. Reinberg. 2005. Histone variants meet their match. *Nat Rev Mol Cell Biol* **6**: 139-49.

Sato, K., M.D. Simon, A.M. Levin, K.M. Shokat, and G.A. Weiss. 2004. Dissecting the Engrailed homeodomain-DNA interaction by phage-displayed shotgun scanning. *Chem Biol* **11**: 1017-23.

Savkur, R.S. and T.P. Burris. 2004. The coactivator LXXLL nuclear receptor recognition motif. *J Pept Res* **63**: 207-12.

Schaaf, M.J. and J.A. Cidlowski. 2003. Molecular determinants of glucocorticoid receptor mobility in living cells: the importance of ligand affinity. *Mol Cell Biol* **23**: 1922-34.

Schaaf, M.J., L.J. Lewis-Tuffin, and J.A. Cidlowski. 2005. Ligand-selective targeting of the glucocorticoid receptor to nuclear subdomains is associated with decreased receptor mobility. *Mol Endocrinol* **19**: 1501-15.

Schaaf, M.J., L. Willetts, B.P. Hayes, B. Maschera, E. Stylianou, and S.N. Farrow. 2006. The relationship between intranuclear mobility of the NF-kappaB subunit p65 and its DNA binding affinity. *J Biol Chem* **281**: 22409-20.

Schaufele, F., J.F. Enwright, 3rd, X. Wang, C. Teoh, R. Srihari, R. Erickson, O.A. MacDougald, and R.N. Day. 2001. CCAAT/enhancer binding protein alpha assembles essential cooperating factors in common subnuclear domains. *Mol Endocrinol* **15**: 1665-76.

Schedl, A., A. Ross, M. Lee, D. Engelkamp, P. Rashbass, V. van Heyningen, and N.D. Hastie. 1996. Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell* **86**: 71-82.

Schier, A.F. and W.J. Gehring. 1992. Direct homeodomain-DNA interaction in the autoregulation of the fushi tarazu gene. *Nature* **356**: 804-7.

Schneider, R. and R. Grosschedl. 2007. Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Genes Dev* **21**: 3027-43.

Schneitz, K., P. Spielmann, and M. Noll. 1993. Molecular genetics of aristaless, a prd-type homeo box gene involved in the morphogenesis of proximal and distal pattern elements in a subset of appendages in Drosophila. *Genes Dev* **7**: 114-29.

Scholl, F.A., J. Kamarashev, O.V. Murmann, R. Geertsen, R. Dummer, and B.W. Schafer. 2001. PAX3 is expressed in human melanomas and contributes to tumor cell survival. *Cancer Res* **61**: 823-6.

Schotta, G., M. Lachner, K. Sarma, A. Ebert, R. Sengupta, G. Reuter, D. Reinberg, and T. Jenuwein. 2004. A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18**: 1251-62.

Schulte, T.W., J.A. Toretsky, E. Ress, L. Helman, and L.M. Neckers. 1997. Expression of PAX3 in Ewing's sarcoma family of tumors. *Biochem Mol Med* **60**: 121-6.

Scully, K.M., E.M. Jacobson, K. Jepsen, V. Lunyak, H. Viadiu, C. Carriere, D.W. Rose, F. Hooshmand, A.K. Aggarwal, and M.G. Rosenfeld. 2000. Allosteric effects of Pit-1 DNA sites on long-term repression in cell type specification. *Science* **290**: 1127-31.

Seksek, O., J. Biwersi, and A.S. Verkman. 1997. Translational diffusion of macromolecule-sized solutes in cytoplasm and nucleus. *J Cell Biol* **138**: 131-42.

Semina, E.V., R. Reiter, N.J. Leysens, W.L. Alward, K.W. Small, N.A. Datson, J. Siegel-Bartelt, D. Bierke-Nelson, P. Bitoun, B.U. Zabel, J.C. Carey, and J.C. Murray. 1996. Cloning and characterization of a novel bicoid-related homeobox transcription factor gene, RIEG, involved in Rieger syndrome. *Nat Genet* **14**: 392-9.

Shapiro, D.N., J.E. Sublett, B. Li, J.R. Downing, and C.W. Naeve. 1993. Fusion of PAX3 to a member of the forkhead family of transcription factors in human alveolar rhabdomyosarcoma. *Cancer Res* **53**: 5108-12.

Sharp, Z.D., D.L. Stenoien, M.G. Mancini, Ouspenski, II, and M.A. Mancini. 2004. Inactivating Pit-1 mutations alter subnuclear dynamics suggesting a protein misfolding and nuclear stress response. *J Cell Biochem* **92**: 664-78.

Singh, S., C.M. Stellrecht, H.K. Tang, and G.F. Saunders. 2000. Modulation of PAX6 homeodomain function by the paired domain. *J Biol Chem* **275**: 17306-13.

Smit, D.J., A.G. Smith, P.G. Parsons, G.E. Muscat, and R.A. Sturm. 2000. Domains of Brn-2 that mediate homodimerization and interaction with general and melanocytic transcription factors. *Eur J Biochem* **267**: 6413-22.

Smith, S.T. and J.B. Jaynes. 1996. A conserved region of engrailed, shared among all en-, gsc-, Nk1-, Nk2- and msh-class homeoproteins, mediates active transcriptional repression in vivo. *Development* **122**: 3141-50.

Spector, D.L. 1996. Nuclear organization and gene expres[illegible]n. *Exp Cell Res* **229**: 189-97.

-. 2001. Nuclear domains. *J Cell Sci* **114**: 2891-3.

-. 2003. The dynamics of chromosome organization and gene regulation. *Annu Rev Biochem* **72**: 573-608.

Stamataki, D., M. Kastrinaki, B.S. Mankoo, V. Pachnis, and D. Karagogeos. 2001. Homeodomain proteins Mox1 and Mox2 associate with Pax1 and Pax3 transcription factors. *FEBS Lett* **499**: 274-8.

Steel, K.P. and R.J. Smith. 1992. Normal hearing in Splotch (Sp/+), the mouse homologue of Waardenburg syndrome type 1. *Nat Genet* **2**: 75-9.

Stenoien, D.L., A.C. Nye, M.G. Mancini, K. Patel, M. Dutertre, B.W. O'Malley, C.L. Smith, A.S. Belmont, and M.A. Mancini. 2001. Ligand-mediated assembly and real-time cellular dynamics of estrogen receptor alpha-coactivator complexes in living cells. *Mol Cell Biol* **21**: 4404-12.

Stifani, S., C.M. Blaumueller, N.J. Redhead, R.E. Hill, and S. Artavanis-Tsakonas. 1992. Human homologs of a Drosophila Enhancer of split gene product define a novel family of nuclear proteins. *Nat Genet* **2**: 119-27.

Strahl, B.D. and C.D. Allis. 2000. The language of covalent histone modifications. *Nature* **403**: 41-5.

Stuart, E.T., C. Kioussi, and P. Gruss. 1994. Mammalian Pax genes. *Annu Rev Genet* **28**: 219-36.

Tajbakhsh, J., H. Luz, H. Bornfleth, S. Lampel, C. Cremer, and P. Lichter. 2000. Spatial distribution of GC- and AT-rich DNA sequences within human chromosome territories. *Exp Cell Res* **255**: 229-37.

Tajbakhsh, S., D. Rocancourt, G. Cossu, and M. Buckingham. 1997. Redefining the genetic hierarchies controlling skeletal myogenesis: Pax-3 and Myf-5 act upstream of MyoD. *Cell* **89**: 127-38.

Tang, Q.Q. and M.D. Lane. 1999. Activation and centromeric localization of CCAAT/enhancer-binding proteins during the mitotic clonal expansion of adipocyte differentiation. *Genes Dev* **13**: 2231-41.

Tang, Y., R.H. Getzenberg, B.N. Vietmeier, M.R. Stallcup, M. Eggert, R. Renkawitz, and D.B. DeFranco. 1998. The DNA-binding and tau2 transactivation domains of the rat glucocorticoid receptor constitute a nuclear matrix-targeting signal. *Mol Endocrinol* **12**: 1420-31.

Tassabehji, M., A.P. Read, V.E. Newton, R. Harris, R. Balling, P. Gruss, and T. Strachan. 1992. Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355**: 635-6.

Tassabehji, M., A.P. Read, V.E. Newton, M. Patton, P. Gruss, R. Harris, and T. Strachan. 1993. Mutations in the PAX3 gene causing Waardenburg syndrome type 1 and type 2. *Nat Genet* **3**: 26-30.

Tazawa, H., W. Osman, Y. Shoji, E. Treuter, J.A. Gustafsson, and J. Zilliacus. 2003. Regulation of subnuclear localization is associated with a mechanism for nuclear receptor corepression by RIP140. *Mol Cell Biol* **23**: 4187-98.

Thomas, J.O. 1999. Histone H1: location and role. *Curr Opin Cell Biol* **11**: 312-7.

Tomilin, A., A. Remenyi, K. Lins, H. Bak, S. Leidel, G. Vriend, M. Wilmanns, and H.R. Scholer. 2000. Synergism with the coactivator OBF-1 (OCA-B, BOB-1) is mediated by a specific POU dimer configuration. *Cell* **103**: 853-64.

Tompa, P. and M. Fuxreiter. 2008. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci* **33**: 2-8.

Ton, C.C., H. Hirvonen, H. Miwa, M.M. Weil, P. Monaghan, T. Jordan, V. van Heyningen, N.D. Hastie, H. Meijers-Heijboer, M. Drechsler, and et al. 1991. Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. *Cell* **67**: 1059-74.

Treisman, J., P. Gonczy, M. Vashishtha, E. Harris, and C. Desplan. 1989. A single amino acid can determine the DNA binding specificity of homeodomain proteins. *Cell* **59**: 553-62.

Treisman, J., E. Harris, and C. Desplan. 1991. The paired box encodes a second DNA-binding domain in the paired homeo domain protein. *Genes Dev* **5**: 594-604.

Tremblay, P. and P. Gruss. 1994. Pax: genes for mice and men. *Pharmacol Ther* **61**: 205-26.

Tremblay, P., F. Pituello, and P. Gruss. 1996. Inhibition of floor plate differentiation by Pax3: evidence from ectopic expression in transgenic mice. *Development* **122**: 2555-67.

Tsukamoto, K., Y. Nakamura, and N. Niikawa. 1994. Isolation of two isoforms of the PAX3 gene transcripts and their tissue-specific alternative expression in human adult tissues. *Hum Genet* **93**: 270-4.

Tsukamoto, K., T. Tohma, T. Ohta, K. Yamakawa, Y. Fukushima, Y. Nakamura, and N. Niikawa. 1992. Cloning and characterization of the inversion breakpoint at chromosome 2q35 in a patient with Waardenburg syndrome type I. *Hum Mol Genet* **1**: 315-7.

Tucker-Kellogg, L., M.A. Rould, K.A. Chambers, S.E. Ades, R.T. Sauer, and C.O. Pabo. 1997. Engrailed (Gln50-->Lys) homeodomain-DNA complex at 1.9 A resolution: structural basis for enhanced affinity and altered specificity. *Structure* **5**: 1047-54.

Tumbar, T. and A.S. Belmont. 2001. Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. *Nat Cell Biol* **3**: 134-9.

Tumbar, T., G. Sudlow, and A.S. Belmont. 1999. Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. *J Cell Biol* **145**: 1341-54.

Underhill, D.A. 2000. Genetic and biochemical diversity in the Pax gene family. *Biochem Cell Biol* **78**: 629-38.

Underhill, D.A. and P. Gros. 1997. The paired-domain regulates DNA binding by the homeodomain within the intact Pax-3 protein. *J Biol Chem* **272**: 14175-82.

Underhill, D.A., K.J. Vogan, and P. Gros. 1995. Analysis of the mouse Splotch-delayed mutation indicates that the Pax-3 paired domain can influence homeodomain DNA-binding activity. *Proc Natl Acad Sci U S A* **92**: 3692-6.

Valenta, T., J. Lukas, L. Doubravska, B. Fafilek, and V. Korinek. 2006. HIC1 attenuates Wnt signaling by recruitment of TCF-4 and beta-catenin to the nuclear bodies. *Embo J* **25**: 2326-37.

Valentine, S.A., G. Chen, T. Shandala, J. Fernandez, S. Mische, R. Saint, and A.J. Courey. 1998. Dorsal-mediated repression requires the formation of a multiprotein repression complex at the ventral silencer. *Mol Cell Biol* **18**: 6584-94.

van Attikum, H. and S.M. Gasser. 2005. The histone code at DNA breaks: a guide to repair? *Nat Rev Mol Cell Biol* **6**: 757-65.

van Driel, R., P.F. Fransz, and P.J. Verschure. 2003. The eukaryotic genome: a system regulated at different hierarchical levels. *J Cell Sci* **116**: 4067-75.

van Holde, K. and J. Zlatanova. 2006. Scanning chromatin: a new paradigm? *J Biol Chem* **281**: 12197-200.

van Steensel, B., M. Brink, K. van der Meulen, E.P. van Binnendijk, D.G. Wansink, L. de Jong, E.R. de Kloet, and R. van Driel. 1995. Localization of the glucocorticoid receptor in discrete clusters in the cell nucleus. *J Cell Sci* **108 (Pt 9)**: 3003-11.

van Steensel, B., E.P. van Binnendijk, C.D. Hornsby, H.T. van der Voort, Z.S. Krozowski, E.R. de Kloet, and R. van Driel. 1996. Partial colocalization of glucocorticoid and mineralocorticoid receptors in discrete compartments in nuclei of rat hippocampus neurons. *J Cell Sci* **109 (Pt 4)**: 787-92.

Vazquez, J., A.S. Belmont, and J.W. Sedat. 2001. Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus. *Curr Biol* **11**: 1227-39.

Verastegui, C., K. Bille, J.P. Ortonne, and R. Ballotti. 2000. Regulation of the microphthalmia-associated transcription factor gene by the Waardenburg syndrome type 4 gene, SOX10. *J Biol Chem* **275**: 30757-60.

Verschure, P.J., I. Van Der Kraan, J.M. Enserink, M.J. Mone, E.M. Manders, and R. Van Driel. 2002. Large-scale chromatin organization and the localization of proteins involved in gene expression in human cells. *J Histochem Cytochem* **50**: 1303-12.

Verschure, P.J., I. van Der Kraan, E.M. Manders, and R. van Driel. 1999. Spatial relationship between transcription sites and chromosome territories. *J Cell Biol* **147**: 13-24.

Visser, A.E., R. Eils, A. Jauch, G. Little, P.J. Bakker, T. Cremer, and J.A. Aten. 1998. Spatial distributions of early and late replicating chromatin in interphase chromosome territories. *Exp Cell Res* **243**: 398-407.

Vogan, K.J., D.J. Epstein, D.G. Trasler, and P. Gros. 1993. The splotch-delayed (Spd) mouse mutant carries a point mutation within the paired box of the Pax-3 gene. *Genomics* **17**: 364-9.

Vogan, K.J. and P. Gros. 1997. The C-terminal subdomain makes an important contribution to the DNA binding activity of the Pax-3 paired domain. *J Biol Chem* **272**: 28289-95.

Vogan, K.J., D.A. Underhill, and P. Gros. 1996. An alternative splicing event in the Pax-3 paired domain identifies the linker region as a key determinant of paired domain DNA-binding activity. *Mol Cell Biol* **16**: 6677-86.

Volpi, E.V., E. Chevret, T. Jones, R. Vatcheva, J. Williamson, S. Beck, R.D. Campbell, M. Goldsworthy, S.H. Powis, J. Ragoussis, J. Trowsdale, and D. Sheer. 2000. Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *J Cell Sci* **113 (Pt 9)**: 1565-76.

Voss, T.C., I.A. Demarco, C.F. Booker, and R.N. Day. 2005. Functional interactions with Pit-1 reorganize co-repressor complexes in the living cell nucleus. *J Cell Sci* **118**: 3277-88.

Vradii, D., S.K. Zaidi, J.B. Lian, A.J. van Wijnen, J.L. Stein, and G.S. Stein. 2005. Point mutation in AML1 disrupts subnuclear targeting, prevents myeloid differentiation, and effects a transformation-like phenotype. *Proc Natl Acad Sci U S A* **102**: 7174-9.

Wachtel, M., M. Dettling, E. Koscielniak, S. Stegmaier, J. Treuner, K. Simon-Klingenstein, P. Buhlmann, F.K. Niggli, and B.W. Schafer. 2004. Gene expression signatures identify rhabdomyosarcoma subtypes and detect a novel t(2;2)(q35;p23) translocation fusing PAX3 to NCOA1. *Cancer Res* **64**: 5539-45.

Wada, H., P.W. Holland, S. Sato, H. Yamamoto, and N. Satoh. 1997. Neural tube is partially dorsalized by overexpression of HrPax-37: the ascidian homologue of Pax-3 and Pax-7. *Dev Biol* **187**: 240-52.

Walther, C., J.L. Guenet, D. Simon, U. Deutsch, B. Jostes, M.D. Goulding, D. Plachov, R. Balling, and P. Gruss. 1991. Pax: a murine multigene family of paired box-containing genes. *Genomics* **11**: 424-34.

Wang, Q., S. Kumar, M. Slevin, and P. Kumar. 2006. Functional Analysis of Alternative Isoforms of the Transcription Factor PAX3 in Melanocytes In vitro. *Cancer Res* **66**: 8574-80.

Wansink, D.G., W. Schul, I. van der Kraan, B. van Steensel, R. van Driel, and L. de Jong. 1993. Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J Cell Biol* **122**: 283-93.

Watanabe, A., K. Takeda, B. Ploplis, and M. Tachibana. 1998. Epistatic relationship between Waardenburg syndrome genes MITF and PAX3. *Nat Genet* **18**: 283-6.

Weber, M. and D. Schubeler. 2007. Genomic patterns of DNA methylation: targets and function of an epigenetic mark. *Curr Opin Cell Biol* **19**: 273-80.

Wheat, W., D. Fitzsimmons, H. Lennox, S.R. Krautkramer, L.N. Gentile, L.P. McIntosh, and J. Hagman. 1999. The highly conserved beta-hairpin of the paired DNA-binding domain is required for assembly of Pax-Ets ternary complexes. *Mol Cell Biol* **19**: 2231-41.

Wiblin, A.E., W. Cui, A.J. Clark, and W.A. Bickmore. 2005. Distinctive nuclear organisation of centromeres and regions involved in pluripotency in human embryonic stem cells. *J Cell Sci* **118**: 3861-8.

Wiggan, O., A. Taniguchi-Sidle, and P.A. Hamel. 1998. Interaction of the pRB-family proteins with factors containing paired-like homeodomains. *Oncogene* **16**: 227-36.

Williams, R.R., V. Azuara, P. Perry, S. Sauer, M. Dvorkina, H. Jorgensen, J. Roix, P. McQueen, T. Misteli, M. Merkenschlager, and A.G. Fisher. 2006. Neural induction promotes large-scale chromatin reorganisation of the Mash1 locus. *J Cell Sci* **119**: 132-40.

Williams, R.R., S. Broad, D. Sheer, and J. Ragoussis. 2002. Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interphase nuclei. *Exp Cell Res* **272**: 163-75.

Wilson, D., G. Sheng, T. Lecuit, N. Dostatni, and C. Desplan. 1993. Cooperative dimerization of paired class homeo domains on DNA. *Genes Dev* **7**: 2120-34.

Xiao, T., H. Hall, K.O. Kizer, Y. Shibata, M.C. Hall, C.H. Borchers, and B.D. Strahl. 2003. Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev* **17**: 654-63.

Xing, H., D.C. Wilkerson, C.N. Mayhew, E.J. Lubert, H.S. Skaggs, M.L. Goodson, Y. Hong, O.K. Park-Sarge, and K.D. Sarge. 2005. Mechanism of hsp70i gene bookmarking. *Science* **307**: 421-3.

Xu, H.E., M.A. Rould, W. Xu, J.A. Epstein, R.L. Maas, and C.O. Pabo. 1999. Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes Dev* **13**: 1263-75.

Xu, W., M.A. Rould, S. Jun, C. Desplan, and C.O. Pabo. 1995. Crystal structure of a paired domain-DNA complex at 2.5 A resolution reveals structural basis for Pax developmental mutations. *Cell* **80**: 639-50.

Yang, X.M., K. Vogan, P. Gros, and M. Park. 1996. Expression of the met receptor tyrosine kinase in muscle progenitor cells in somites and limbs is absent in Splotch mice. *Development* **122**: 2163-71.

Yavuzer, U. and C.R. Goding. 1994. Melanocyte-specific gene expression: role of repression and identification of a melanocyte-specific factor, MSF. *Mol Cell Biol* **14**: 3494-503.

Yin, X., M.F. Landay, W. Han, E.S. Levitan, S.C. Watkins, R.M. Levenson, D.L. Farkas, and E.V. Prochownik. 2001. Dynamic in vivo interactions among Myc network members. *Oncogene* **20**: 4650-64.

Young, D.W., M.Q. Hassan, X.Q. Yang, M. Galindo, A. Javed, S.K. Zaidi, P. Furcinitti, D. Lapointe, M. Montecino, J.B. Lian, J.L. Stein, A.J. van Wijnen, and G.S. Stein. 2007. Mitotic retention of gene expression patterns by the cell fate-determining transcription factor Runx2. *Proc Natl Acad Sci U S A* **104**: 3189-94.

Zaidi, S.K., A. Javed, J.Y. Choi, A.J. van Wijnen, J.L. Stein, J.B. Lian, and G.S. Stein. 2001. A specific targeting signal directs Runx2/Cbfa1 to subnuclear domains and contributes to transactivation of the osteocalcin gene. *J Cell Sci* **114**: 3093-102.

Zaidi, S.K., A.J. Sullivan, R. Medina, Y. Ito, A.J. Van Wijnen, J.L. Stein, J.B. Lian, and G.S. Stein. 2004. Tyrosine phosphorylation controls Runx2-mediated subnuclear targeting of YAP to repress transcription. *Embo J* **23**: 790-9.

Zaidi, S.K., A.J. Sullivan, A.J. van Wijnen, J.L. Stein, G.S. Stein, and J.B. Lian. 2002. Integration of Runx and Smad regulatory signals at transcriptionally active subnuclear sites. *Proc Natl Acad Sci U S A* **99**: 8048-53.

Zeng, C., S. McNeil, S. Pockwinse, J. Nickerson, L. Shopland, J.B. Lawrence, S. Penman, S. Hiebert, J.B. Lian, A.J. van Wijnen, J.L. Stein, and G.S. Stein. 1998. Intranuclear targeting of AML/CBFalpha regulatory factors to nuclear matrix-associated transcriptional domains. *Proc Natl Acad Sci U S A* **95**: 1585-9.

Zeng, C., A.J. van Wijnen, J.L. Stein, S. Meyers, W. Sun, L. Shopland, J.B. Lawrence, S. Penman, J.B. Lian, G.S. Stein, and S.W. Hiebert. 1997. Identification of a nuclear matrix targeting signal in the leukemia and bone-related AML/CBF-alpha transcription factors. *Proc Natl Acad Sci U S A* **94**: 6746-51.

Zhang, L. and C. Wang. 2007. Identification of a new class of PAX3-FKHR target promoters: a role of the Pax3 paired box DNA binding domain. *Oncogene* **26**: 1595-605.

Zhang, R., M.V. Poustovoitov, X. Ye, H.A. Santos, W. Chen, S.M. Daganzo, J.P. Erzberger, I.G. Serebriiskii, A.A. Canutescu, R.L. Dunbrack, J.R. Pehrson, J.M. Berger, P.D. Kaufman, and P.D. Adams. 2005. Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. *Dev Cell* **8**: 19-30.

Zink, D., M.D. Amaral, A. Englmann, S. Lang, L.A. Clarke, C. Rudolph, F. Alt, K. Luther, C. Braz, N. Sadoni, J. Rosenecker, and D. Schindelhauer. 2004. Transcription-dependent spatial arrangements of CFTR and adjacent genes in human cell nuclei. *J Cell Biol* **166**: 815-25.

Zirbel, R.M., U.R. Mathieu, A. Kurz, T. Cremer, and P. Lichter. 1993. Evidence for a nuclear compartment of transcription and splicing located at chromosome domain boundaries. *Chromosome Res* **1**: 93-106.

APPENDIX 1. EFFECT OF THE INTERDOMAIN LINKER ON PAX3 DNA BINDING

A1.1 Role of the interdomain linker in PAX3 homeodomain DNA binding

The optimal binding motif for the PAX3 HD is the *P2* sequence, which features an inverted palindrome composed of 5'-TAAT half-sites separated by 2 bp (Wilson et al. 1993). We showed in Chapter 3 that the HD can also interact with a variety of AT-rich motifs, and that HD binding is an important aspect of composite sequence recognition by PAX3. While our prior results demonstrated that tethering the HD to the PD improves PAX3 affinity for composite promoter sequences (Fig. 3-4), they did not provide any information on the effect of the interdomain linker during this activity. Since previous analyses of PAX3 DNA binding suggested portions of the HD amino-terminal extension and interdomain linker influenced HD-DNA interaction and PD-HD interdependence (Fortin et al. 1998), we undertook a detailed investigation of the linker's effects on PAX3 HD DNA binding (below) and cooperativity between the PD and HD (Section A1.2).

To examine the effect of the linker on HD DNA binding activity, we created recombinant deletion constructs containing the PAX3 HD and various portions of linker (Fig. A1-1A), which allowed us to assess the contribution of the octapeptide, NLS, amino-terminal extension, and recognition helix on HD behavior. We initially tested binding to the *in vitro*-optimized *P1/2* and *P2* sequences, which support monomeric and dimeric binding, respectively. The HD+linker construct, which contains the full linker, bound only to the *P2* sequence (Fig. A1-1B, *top panel*), suggesting cooperative dimerization is necessary for HD+linker binding. This construct also bound to the *P3* sequence (not shown), which, like *P2*, supports cooperative HD binding but features palindromic TAAT half sites separated by 3 bp (Wilson et al. 1993). Gradual deletion of the amino-terminal portion of the linker permitted binding of the HD+Oct, HD-Oct,

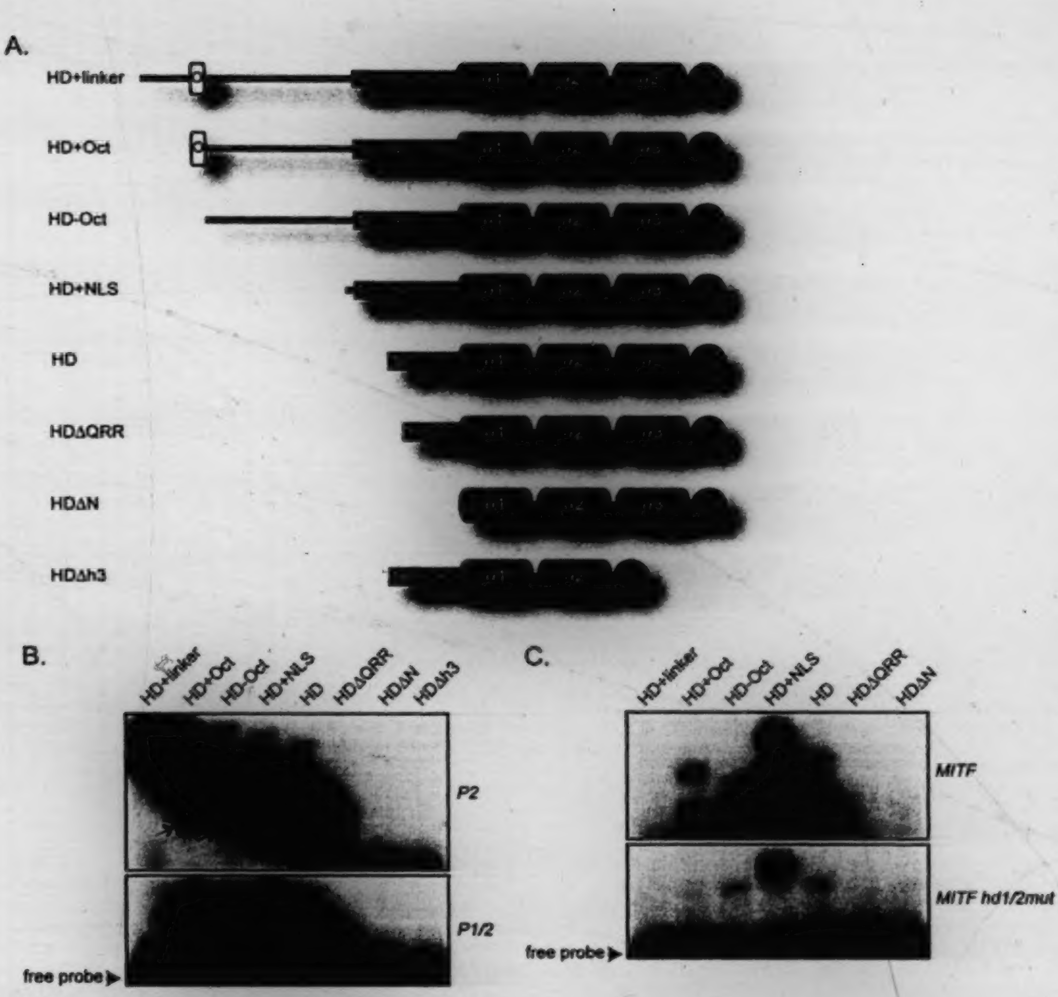

A.
HD+linker
HD+Oct
HD-Oct
HD+NLS
HD
HDΔQRR
HDΔN
HDΔh3
B.
HD+linker
HD+Oct
HD-Oct
HD+NLS
HD
HDΔQRR
HDΔN
HDΔh3
P2
P1/2
free probe
C.
HD+linker
HD+Oct
HD-Oct
HD+NLS
HD
HDΔQRR
HDΔN
MITF
MITF hd1/2mut
free probe

Figure A1-1. The PAX3 interdomain linker influences homeodomain DNA binding behavior. (A) Recombinant peptides encoding the PAX3 HD and various portions of the interdomain linker or deletions of the HD were created to address the role of the linker and determinants of HD DNA binding. O, octapeptide; NLS, nuclear localization signal; N-term arm, amino-terminal HD extension; His, hexahistidine tag. See Section 2.1.1 for details. (B) Mobility shift assays using the HD constructs and *P2* (*top*) or *P1/2* (*bottom*) oligonucleotides were used to assess influence of the linker and HD regions on DNA binding and dimerization. Black arrows indicate monomeric binding to *P2*. (C) Mobility shift assays using the *MITF* (*top*) and *MITF hd1/2mut* (*bottom*) oligonucleotides allowed investigation of the role of the linker and HD regions in binding to suboptimal recognition sequences. Position of free probe is indicated for bottom panels only.

HD+NLS, and HD proteins to both oligonucleotides – each protein dimerized on *P2*, with various degrees of residual monomeric binding (Fig. A1-1B, *arrows*). Interestingly, the HD-Oct and HD+NLS constructs each bound to the *P1/2* site in two distinct complexes (Fig. A1-1B, *bottom panel*), as indicated by bands of different mobilities. Deletion of the first three HD residues (HDΔQRR) abrogated binding to both sequences, although this construct binds to DNA at sufficiently high concentrations (not shown). Deletion of the amino-terminal extension (HDΔN) or recognition helix (HDΔh3) also precluded binding, consistent with studies showing that these regions harbor residues involved in DNA contact (Gehring et al. 1994b).

With the exception of HDΔh3, we also tested binding of the HD constructs to the *MITF* and *MITF hd1/2mut* oligonucleotides, which allowed us to evaluate interaction with suboptimal HD sites (Fig. A1-1C). We previously showed that the HD can interact with the suboptimal hd3 motif of the *MITF* sequence, even though the hd1/2 motifs are the predominant sites of binding (Fig. 3-6), and that the *MITF* sequence supports both monomeric and dimeric HD binding (Fig. 3-8A). We observed no binding of HD+linker to either sequence, again indicating the requirement not only for a cooperative binding event, but also for a sequence containing properly spaced TAAT half-sites (the *MITF* sequence contains multiple HD-binding motifs, but none with the spacing/orientation of *P2* or *P3*). The HD+Oct protein displayed low affinity monomeric and dimeric binding to *MITF* but did not bind to *MITF hd1/2mut*. In contrast, the HD-Oct protein, while failing to interact with the *MITF hd1/2mut* sequence, bound to *MITF* with moderate affinity, suggesting that the octapeptide might affect HD interaction with certain recognition elements. The HD+NLS protein bound to *MITF* with high affinity in both monomeric and dimeric forms and, surprisingly, displayed monomeric binding to *MITF hd1/2mut*. As before, the HD construct bound strongly as a monomer and with less affinity as a dimer to the *MITF* sequence, but we observed no interaction with the *MITF hd1/2mut* oligonucleotide. Together, these results suggest that the NLS and/or adjacent residues improve affinity of the HD for DNA, and enhance HD interaction with suboptimal recognition sites. This may be due to an overall improvement in the stability of the HD-DNA complex or the influence of charged residues that are present in the HD+NLS protein. As expected, neither the HDΔQRR nor HDΔN constructs interacted with the two oligonucleotides, again demonstrating the importance of the amino-terminal extension in HD DNA binding.

A1.2 Influence of the linker on cooperative binding

The PAX3 interdomain linker, like the linker separating the POU subdomains of OCT1 (Klemm et al. 1994), is likely unstructured in solution; however, it features several distinct regions that may harbor determinants of PAX3 activity. The most obvious of these is the octapeptide which, based on its similarity to Eh-1-like motifs, is predicted to form a short amphipathic helix (Savkur and Burris 2004; Copley 2005). The linker is also highly charged, with half of its 56 residues being acidic or basic amino acids (Fig. A1-2A). Having shown that the presence or absence of various portions of the linker affects HD DNA binding affinity (above), we performed recruitment binding assays with PD or HD derivatives containing portions of the linker to assess its effects on PD-HD cooperativity. The PD+linker protein, which contains the full linker fused to the PD, was more effective than the PD alone at recruiting the HD or HD+NLS proteins to

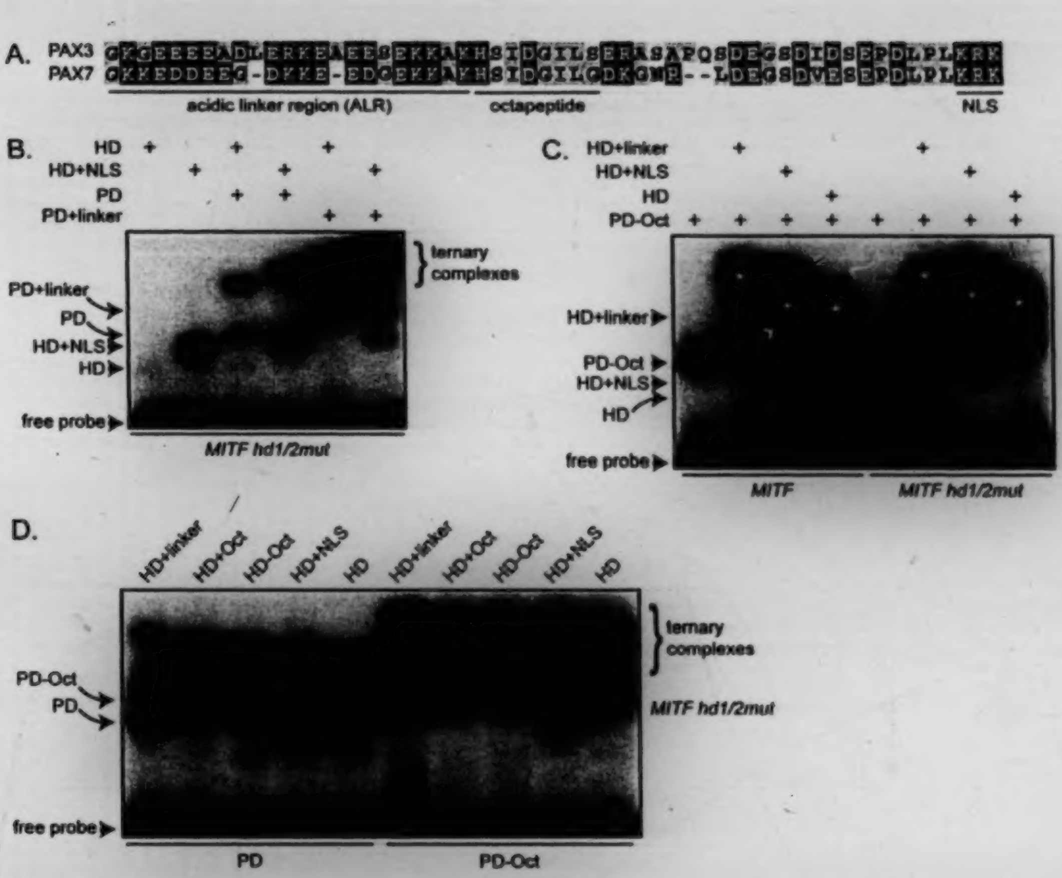

A.
PAX3 GKGEEEEADLERKEAEESEKKAKHSIDGILSERASAPQSDEGSDIDSEPDLPLKRK
PAX7 GKKEDDEEG-DKKE-EDGEKKAKHSIDGILGDKGNR--LDEGSDVESEPDLPLKRK
acidic linker region (ALR)
octapeptide
NLS
B.
HD
HD+NLS
PD
PD+linker
ternary complexes
PD+linker
PD
HD+NLS
HD
free probe
MITF hd1/2mut
C.
HD+linker
HD+NLS
HD
PD-Oct
HD+linker
PD-Oct
HD+NLS
HD
free probe
MITF
MITF hd1/2mut
D.
HD+linker
HD+Oct
HD-Oct
HD+NLS
HD
HD+linker
HD+Oct
HD-Oct
HD+NLS
HD
ternary complexes
MITF hd1/2mut
PD-Oct
PD
free probe
PD
PD-Oct

Figure A1-2. The interdomain linker affects cooperative binding by the paired domain and homeodomain. (A) Sequence alignment of the PAX3 and PAX7 interdomain linkers (sequences correspond to aa 163-218 in PAX3 and aa 163-214 in PAX7). Residues are color-coded as follows: hydrophobic (grey), polar uncharged (yellow), acidic (red), basic (blue). Positions of the acidic linker region (ALR), octapeptide, and nuclear localization signal (NLS) are indicated. (B) Recruitment binding assays were used to assess the role of the linker in cooperative binding by the PAX3 PD and HD. (C) Role of linker segments on cooperative PD-HD binding to the *MITF* (lanes 1-4) and *MITF hd1/2mut* (lanes 5-8) oligonucleotides was examined using recruitment binding assays. Asterisks denote ternary PD-HD-DNA complexes; white arrow, HD+NLS dimer; black arrow, HD+NLS dimer + PD. (D) Binding assays used to examine the influence of the ALR on recruitment of HD constructs. Positions of protein-DNA complexes and free probe are indicated; presence of each protein is indicated above each lane.

MITF hd1/2mut site (Fig. A1-2B). However, PD+linker binding affinity was noticeably higher than that of the PD, implying that increased ternary complex in the presence of PD+linker could merely be due to improved PD-DNA interaction. Interestingly, the PD and PD+linker bands appeared fainter in the presence of the HD+NLS, suggesting the HD+NLS is more efficient at ternary complex formation than the HD alone. As previously demonstrated (Fortin et al. 1998), these results show that the linker plays an important role in cooperative binding by the PAX3 DNA-binding domains.

In both PAX3 and PAX7, the portion of the linker immediately following the PD is rich in acidic and basic residues with a net negative charge (Fig. A1-2A). Having demonstrated that addition of the linker to the PD results in enhancement of PD DNA binding activity (see above), we next assessed the effects of this acidic linker region (ALR) on PD behavior. For this, we created a protein (PD-Oct) containing the PD fused to the ALR and performed recruitment binding assays with selected PAX3 HD derivative proteins. PD-Oct bound with high affinity to both the *MITF* and *MITF hd1/2mut* oligonucleotides and showed efficient cooperative binding in the presence of the HD+linker, HD+NLS, and HD proteins (Fig. A1-2C). Remarkably, HD+NLS formed four unique complexes on *MITF* sequence – monomeric, dimeric (*white arrow*), ternary (HD+NLS + PD-Oct; *asterisk*), and quaternary (2x HD+NLS + PD-Oct; *black arrow*). The HD protein also formed a quaternary complex with PD-Oct on *MITF*, albeit with reduced efficiency compared to HD+NLS. On the *MITF hd1/2mut* oligonucleotide, ternary complex formation appeared more efficient, given the reduction in monomeric binding by HD and HD+NLS. This suggests that the efficiency of the PD in recruiting HD proteins to a composite binding sequence varies depending on the amount of potential HD-binding sites, which would otherwise be occupied by the HD in the absence of the PD.

Finally, we compared the ability of the PD and PD-Oct proteins to form ternary complexes with selected HD deletion constructs. In general, PD-Oct possessed higher affinity for the *MITF hd1/2mut* sequence than the PD and displayed more efficient cooperative binding, as implied by an increase in ternary complex band intensity (Fig. A1-2D). Together with the above data, these results show that addition of linker sequence to the PD improves DNA binding and enhanced cooperativity. The ALR, rather than the full linker, appears sufficient for these alterations. The additional negative charge of this region probably does not directly improve binding affinity, since a charge-repulsion situation would likely occur with the negatively charged DNA. Instead, the ALR may stabilize PD conformation on the DNA through an intramolecular interaction, and may also enhance the stability of PD-HD interaction, either on or off the DNA. The improved affinity and cooperativity of the HD+NLS protein could also be due to charge-related effects, as the NLS features several positively charged residues that may enhance interaction with the negatively charged DNA and/or acidic residues in the ALR.

A1.3 Discussion

Together, these results suggest that the PAX3 interdomain linker wields a certain degree of influence over the activity of the HD and demonstrate that the ALR can influence the DNA binding properties of the PAX3 PD and HD. The presence of this region had a negative effect on HD binding that was alleviated when another copy was provided in *trans*. In particular, interaction of the HD+linker protein with HD-specific sites was only possible in the presence of the *P2* and *P3* sequences, each of which supports cooperative binding of two HD molecules (Fig. A1-1B and data not shown), while ternary complex formation involving HD+linker was enhanced when the PD-Oct protein, which features the ALR, was used (Fig. A1-2D). Our *in vivo* observations of PAX3 deletion constructs also support a role for the linker in modulating HD activity. Notably, removal of the ALR exerted different effects on the subnuclear localization of the HD (Fig. 4-6). Combining this result with the above DNA binding data provides solid evidence that the interdomain linker influences HD behavior and that the ALR plays a prominent role in this activity. Significantly, our results suggest that the ALR might influence PAX3 behavior by exerting an autoregulatory or inhibitory effect on DNA binding activity and that a conformational rearrangement of PAX3 upon DNA binding alleviates this effect. The possible mechanisms involved in such an event are further discussed in Chapter 6.

APPENDIX 2. RELATIONSHIP BETWEEN PAX3 AND THE TLE4 CO-REPRESSOR

A2.1 TLE4: gene and protein structure

PAX3 was recently shown to interact with the GRG4 co-repressor, leading to repression of *Dct*, a marker of melanocyte differentiation (Lang et al. 2005). The GRG/TLE proteins are mammalian homologs of *Drosophila* Gro, a co-repressor of the hairy repressor and hairy-related proteins that function in neuronal differentiation, segmentation, sex determination, and myogenesis (Paroush et al. 1994). GRG/TLE co-repressors interact with DNA-binding proteins to control gene expression during developmental processes (Fisher and Caudy 1998; Chen and Courey 2000) and participate in the regulation of a number of developmentally important pathways (Gasperowicz and Otto 2005; Buscarlet and Stifani 2007). *Grg/TLE* genes exhibit overlapping expression patterns with developmentally important transcription factors, and several different classes of mammalian transcription factors are known to recruit GRG/TLE co-repressors to facilitate repression of target genes (reviewed in Gasperowicz and Otto 2005).

Cloning of the human *TLE4* gene and its murine homolog *Grg4* revealed extensive homology to the *Drosophila Gro* gene and mammalian *TLE/Grg* genes (Stifani et al. 1992; Koop et al. 1996). The human *TLE4* transcript is differentially spliced to produce three distinct proteins (Fig. A2-1A); the canonical isoform includes twenty exons and encodes a 773-aa protein, which hereafter will be referred to as TLE4a (Fig. A2-1A, *top*). A second isoform, designated *TLE4b*, includes all twenty exons plus an additional exon between exons 11 and 12. In the human transcript, this 96 bp insert ("exon 11a") is contained within intron 11 and inclusion of exon 11a inserts an additional 32 aa into the S/P domain, creating a protein of 805 aa. The third isoform, *TLE4c*, lacks exons 10 and

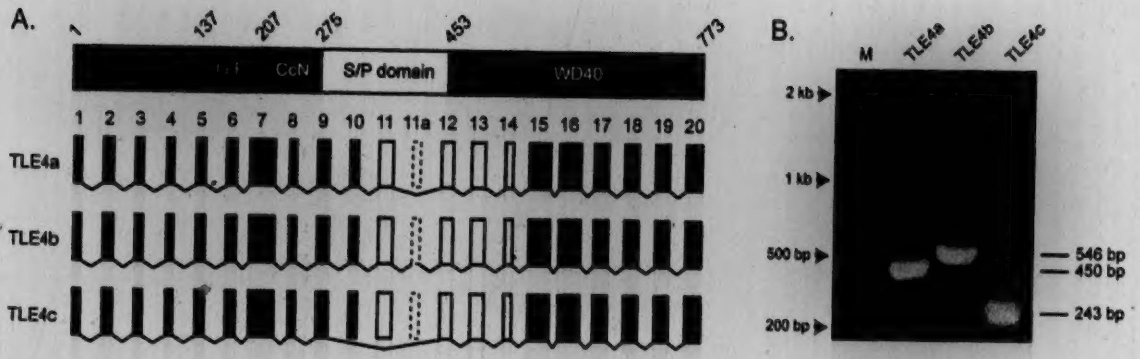

A.
1
137
207
275
453
773
CcN
S/P domain
WD40
1 2 3 4 5 6 7 8 9 10 11 11a 12 13 14 15 16 17 18 19 20
TLE4a
TLE4b
TLE4c
B.
M
TLE4a
TLE4b
TLE4c
2 kb
1 kb
500 bp
200 bp
546 bp
450 bp
243 bp

Figure A2-1. The human *TLE4* transcript is alternatively spliced to produce three isoforms. (A) Schematic diagram of the human TLE4 protein (*top*) and three isoforms (*bottom*) produced by alternative splicing events; see Section A2.1 for details. TLE4 domains are coded by color (Q, Gln-rich domain; G/P, Gly-Pro-rich domain; CcN, CK2 and cdc2/nuclear localization signal domain; S/P, Ser-Pro-rich domain; WD40, Trp-Asp-repeat domain); exons encoding portions of each domain are colored correspondingly. Amino acid positions in the TLE4 protein and exons featured in each transcript are indicated. (B) PCR amplification (see Section 2.1.1 for PCR conditions) of TLE4 fragments using primers extending from the beginning of exon 9 to the end of exon 12 (TLE4.exon9.F; 5'-TATCCCCATCAGCCAGTTTCCGAGGT and TLE4.exon12.R; 5'-CCAAAGGGTCAACTCCTGGTGGTT) was used to verify the existence of three distinct human TLE4 cDNAs. Sizes of each fragment are indicated at left; M, DNA size marker (bp).

11, which code for the end of the CcN domain and the beginning of the S/P domain, and encodes a protein of 704 aa. PCR amplification of the region spanning exons 9 to 12 confirms the existence of three distinct *TLE4* splice variants (Fig. A2-1B). It is currently unknown if the murine *Grg4* transcript is spliced to form the same three isoforms found in humans.

A2.2 Intranuclear relationship between PAX3 and TLE4

Several studies have demonstrated that the TLE proteins, including GRG4/TLE4, localize to the nucleus (Stifani et al. 1992; Milili et al. 2002; Lang et al. 2005). To analyze the localization of endogenous GRG4, we performed immunofluorescence with B16F10 cells, which were shown previously to express GRG4 (Lang et al. 2005). Our results show that GRG4 is excluded from Hoechst-enriched areas, including perinucleolar and pericentromeric regions and the nuclear periphery, and forms multiple bright foci that overlays a diffuse background of lower intensity (Fig. A2-2A). To examine the subnuclear localization of the three human TLE4 isoforms, we expressed each as an amino-terminally HA-tagged fusion protein in 10T1/2-cells. HA-TLE4a formed a reticular pattern that was evenly distributed throughout the nucleus and enriched in perinucleolar regions, but did not overlap with compacted chromatin stained by Hoechst (Fig. A2-2B, *top row*). In contrast, HA-TLE4b formed numerous punctate foci with a low level of diffuse background staining and was excluded from nucleoli and Hoechst-stained chromatin (Fig. A2-2B, *middle row*). While uniform within a single nucleus, focal size differed from cell to cell, within a range of approximately 0.9-1.7 μM. HA-TLE4c also formed numerous foci but, in contrast to HA-TLE4b, we consistently

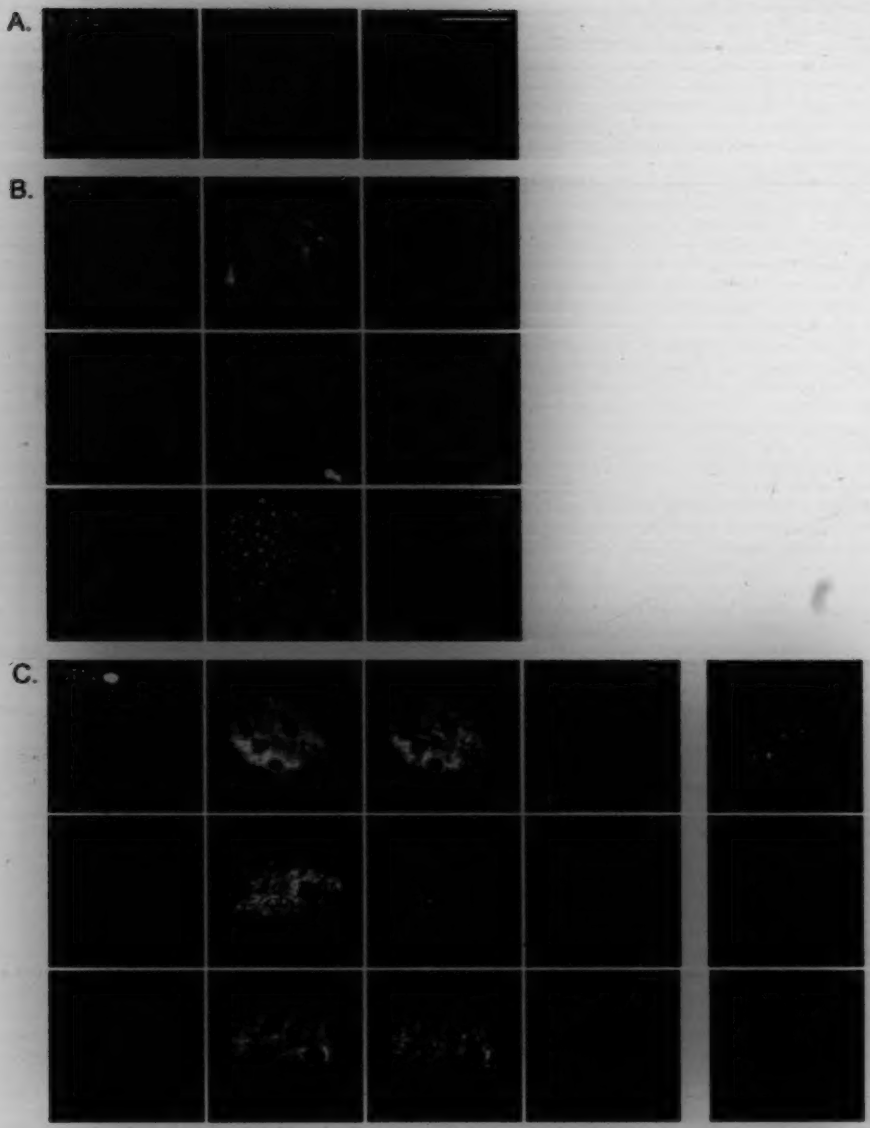


Figure A2-2. Intranuclear relationship between PAX3 and TLE4. (A) Indirect immunofluorescence was used to assess the subnuclear localization of endogenous GRG4 in B16F10 murine melanoma cells. (B) Intranuclear compartmentalization of the TLE4a (*top row*), TLE4b (*middle row*), and TLE4c (*bottom row*) isoforms in 10T1/2 mouse fibroblast cells. Each TLE4 isoform was expressed with an amino-terminal HA epitope tag and an antibody against the HA tag was used for immunofluorescence. (C) 10T1/2 cells were co-transfected with expression plasmids for untagged PAX3 and HA-tagged TLE4a (*top row*), TLE4b (*middle row*), or TLE4c (*bottom row*). Indirect immunofluorescence using antibodies against PAX3 (green) and the HA tag (red) was used to examine the subnuclear distribution of each protein. Signal intensity thresholds (*far right*) detail regions of overlap (yellow) between PAX3 (green) and HA-TLE4 (red). Grayscale panels are unaltered; merged panels have been deconvolved. Bar, 10 μm.

observed more foci per nucleus in HA-TLE4c-transfected cells and non-uniform focal size within a given nucleus (Fig. A2-2B, *bottom row*). GFP-tagged fusions of the TLE4b and TLE4c isoforms showed no difference from HA-tagged versions (not shown).

PAX3 interacts with GRG4 *in vitro* and GRG4 and PAX3 are co-expressed in a subset of hair follicle cells (Lang et al. 2005). At present, however, it is unknown whether PAX3 interacts with any or all of the three human TLE4 isoforms, or if they show any co-localization in cells in which they are co-expressed. We therefore wished to examine the subnuclear organization of PAX3 co-expressed with each TLE4 isoform. We transfected 10T1/2 cells with expression plasmids encoding untagged PAX3 and HA-tagged versions of TLE4a, TLE4b, or TLE4c and performed co-immunofluorescence with antibodies for PAX3 and the HA tag. PAX3 typically showed extensive but incomplete co-localization with HA-TLE4a (Fig. A2-2C, *top row*). Significantly, both HA-TLE4b (Fig. A2-2C, *middle row*) and HA-TLE4c (Fig. A2-2C, *bottom row*) were redistributed from their characteristic foci in the presence of PAX3, despite limited co-localization between PAX3 and these two isoforms. These results suggest that PAX3 can prevent the accumulation of TLE4b and TLE4c in nuclear foci and/or is capable of dissociating and redistributing these TLE4 isoforms to nuclear domains occupied by PAX3.

A2.3 Disease mutations affect PAX3 co-localization with TLE4 isoforms

We have demonstrated that disease mutations affect various aspects of PAX3 function, including DNA binding (Figs. 3-11 & 3-12), reporter gene transactivation (Fig. 3-13), subnuclear compartmentalization (Fig. 4-7), and intranuclear mobility (Fig. 4-8).

To date, however, the effects of PAX3 disease mutations on interactions with other proteins have not been examined. To assess whether PAX3 disease mutations affect its *in vivo* relationship with TLE4, we transfected 10T1/2 cells with each HA-tagged TLE4 isoform and untagged PAX3 featuring each of the ten mutations previously examined for effects on subnuclear localization. Due to the similarity in localization patterns between the wild type and mutant PAX3 variants and HA-TLE4a, no obvious alterations were observed in any of the proteins' intranuclear localization patterns (not shown). However, co-expression of PAX3 disease mutants with the HA-TLE4b or HA-TLE4c revealed varying effects on the distributions of each protein, representatives of which are shown in Figure A2-3. Some mutants, such as F45L, were unable to redistribute HA-TLE4b from its characteristic foci (Fig. A2-3A, *top row*) but displayed no co-localization with TLE4b in foci, and little overlap in the remaining portion of the nucleus. In contrast, neither of the G81A or R271G mutants redistributed HA-TLE4b (Fig. A2-3A, *middle & bottom rows*), although in both cases, a portion of PAX3 co-localized with TLE4b in the foci.

In cells co-expressing HA-TLE4c and PAX3 disease mutants, we observed a similar range of patterns. In the presence of the *Sp*^{d} mutant, TLE4c foci were not dispersed and the majority of co-localization was on the focal periphery (Fig. A2-3B, *top row*). As observed for HA-TLE4b, the G81A and R271G mutants failed to disperse HA-TLE4c from foci, with the majority of co-localization being on the focal periphery (Fig. A2-3B, *second & bottom rows*). A subset of PAX3 mutants, including S84F (Fig. A2-3B, *third row*), redistributed HA-TLE4c from focal domains and we observed moderate overlap between the two signals; however, the TLE4c pattern appeared disorganized with no discernable pattern, suggesting that co-localization in this case is coincidental.

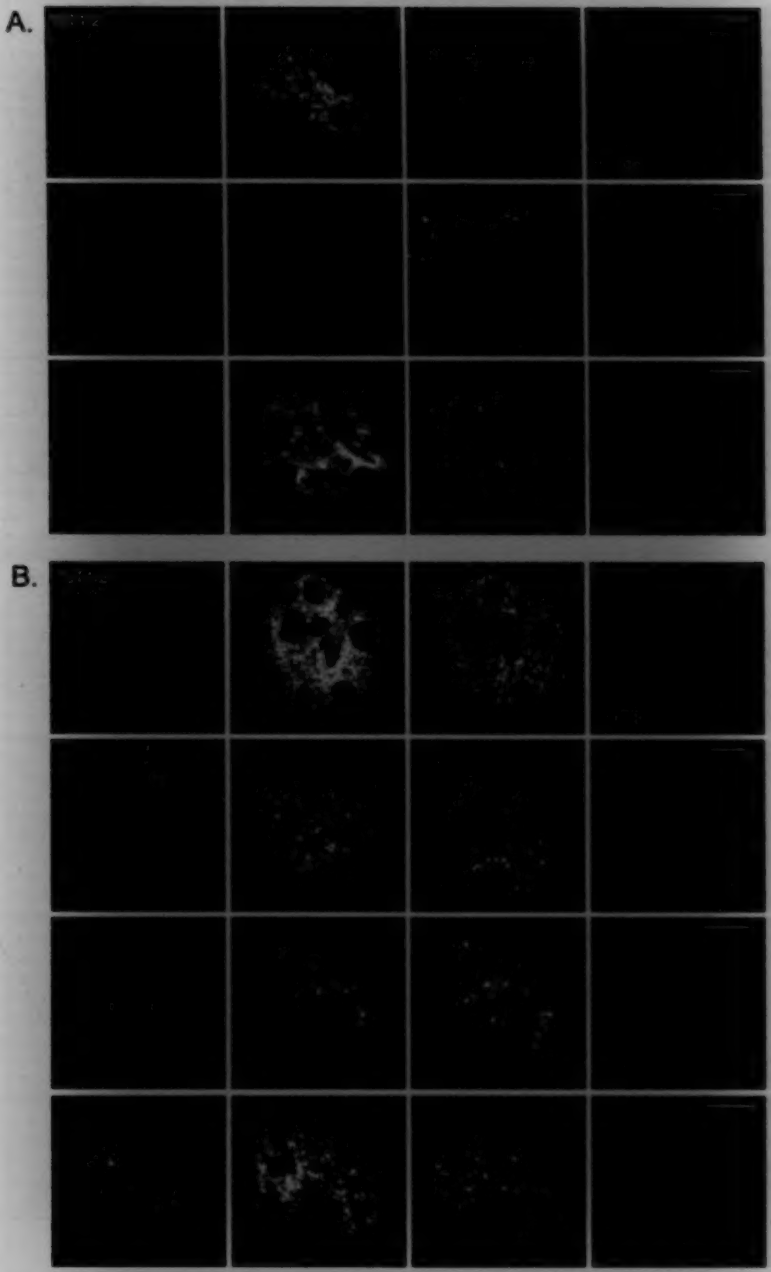


Figure A2-3. Co-localization of TLE4b and TLE4c with selected PAX3 disease mutants. Effects of selected PAX3 disease mutations on the intranuclear distribution of PAX3 and TLE4b (A) or TLE4c (B) were examined by co-transfecting 10T1/2 cells with plasmids encoding untagged PAX3 mutants (green) and HA-tagged TLE4 proteins (red). Grayscale panels are unaltered; merged panels have been deconvolved. Bar, 10 μm.

Together, these preliminary results indicate that disease-causing mutations exert a range of effects on the ability of PAX3 to redistribute TLE4 isoforms from nuclear foci. However, it is still unclear at this point what underlies this mechanism of redistribution and why some mutants, such as G81A and R271G, become mislocalized to TLE4 foci.

A2.4 Discussion

We have demonstrated the existence of three alternatively spliced isoforms of TLE4 that exhibit varying degrees of co-localization with PAX3 in discrete nuclear domains (Fig. A2-2). Although preliminary, our data show that PAX3 disease mutations affect its relationship with TLE4 in the context of the nucleus (Figs. A2-3). These are the first observations to demonstrate that the *in vivo* behavior of a PAX3-interacting protein is altered in the presence of PAX3 disease alleles. Significantly, co-expression of PAX3 mutants and the TLE4b or TLE4c isoforms revealed anomalies in the behavior of each protein. These anomalies can be divided into two categories: (1) the inability of PAX3 to mobilize TLE4 from foci to other regions of the nucleus and (2) the mislocalization of PAX3 into TLE4 foci. In regard to focal localization of PAX3, the G81A and R271G mutants were the only variants found in foci formed by TLE4b and TLE4c. In a similar situation, disease-associated mutations in the AR caused mislocalization to nuclear foci and focal recruitment of AR-interacting proteins (Nazareth et al. 1999; Black et al. 2004). Future investigations will be required to elucidate why only two of the three TLE4 isoforms accumulate in nuclear foci and whether alternative splicing contributes to this behavior. In addition, it will be important to characterize how and why PAX3 and TLE4 are recruited to each other's nuclear domains. Given that GRG4 has been shown to interact with PAX2 (Cai et al. 2003), PAX3 (Lang et al. 2005), and PAX5 (Eberhard et al. 2000), it is possible that the relationship between PAX3 and TLE4 *in vivo* is dependent on physical association, which may be disrupted in the presence of certain disease-causing mutations, contributing to a pathogenic phenotype.